(1)

AFIT/GEE/ENS/93S-03

A PARAMETRIC REGRESSION OF THE COST OF BASE REALIGNMENT ACTION (COBRA) MODEL

THESIS

Douglas D. Hardman, Captain, USAF
Michael S. Nelson, Captain, USAF

AFIT/GEE/ENS/93S-03

186pp

The contents of this document are technically accurate and no sensitive items, detrimental ideas, or deleterious information is contained therein. Furthermore, the views expressed in the document are those of the author(s) and do not necessarily reflect the views of the School of Engineering, the Air University, the Air Education and Training Command, the United States Air Force, or the Department of Defense.

THESIS APPROVAL

STUDENTS: Captain Douglas Hardman
Captain Michael Nelson

CLASS: GEE 93S

TITLE: A Parametric Regression of the Cost of Base Realignment Action (COBRA) Model

DEFENSE DATE: 20 August 1993

COMMITTEE	NAME/DEPARTMENT	SIGNATURE
Co-Advisor	J.P. Cain, PhD Associate Professor of Operations Research Department of Operational Sciences School of Engineering	Joseph P. Cain
Co-Advisor	J.A. Howell, Major, USAF, PhD Assistant Professor of Operations Research Department of Operational Sciences School of Engineering	J.A. Howell
Reader	B.W. Woodruff, Major, USAF, PhD Associate Professor of Statistics Department of Mathematics and Statistics School of Engineering	Brian W. Woodruff

AFIT/GEE/ENS/93S-03

A PARAMETRIC REGRESSION OF THE COST OF BASE REALIGNMENT ACTION (COBRA) MODEL

THESIS

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Engineering

and Environmental Management

Douglas D. Hardman, B.S.E.E.
Captain, USAF

Michael S. Nelson, B.S.C.E
Captain, USAF

September 1993

Approved for public release; distribution unlimited

ACKNOWLEDGMENTS

The authors would like to acknowledge several people who provided invaluable help with this thesis. We would first like to thank Mr. Jeff Miller of the Base Closure Commission for his responsiveness to our questions and requests. Without his help we would probably still be searching for data. Next, we would like to thank Major Charles Fletcher for sending us a copy of the COBRA documentation. Our thanks also goes out to Major Andy Howell and Dr. Joseph Cain for agreeing to be our advisors so late in the thesis process and for pushing us in the right direction toward completion. Professor Dan Reynolds also has made significant contributions to our efforts. His knowledge of regression techniques and SAS capabilities combined with his unending willingness to provide assistance provided us with many answers and immeasurable reassurance.

Finally, and probably most importantly, we would like to thank our wives Captain Amy Hardman and Captain Kristen Nelson. We're grateful for their tolerance, support, and restraint in not pulling rank on us to get home when we were working late into the night.

Table of Contents

Page

List of Figures

Page

Page

Page

List of Tables

AFIT/GEE/ENV/93S-19

Abstract

This study develops a parametric model that is capable of generating accurate estimates of the costs to close Air Force installations. The new model is based upon, but much simpler to use than, the Cost of Base Realignment Action (COBRA) model. COBRA is an economic cost analysis model that requires a minimum of 250 inputs and as many as 700 inputs. The new parametric model requires just 10 input variables and was developed using least squares multiple regression. Comparison of the new parametric model to COBRA indicates that it captures 91% of the variance in cost estimates generated by the detailed COBRA model.

The 20-year Net Present Value (NPV) of actions to close an installation is the figure of merit used in the new model. The COBRA Model and the new parametric model generate similar rank orderings of bases, when NPV is used as the ranking criterion. Use of the Spearman's Correlation Test shows a direct correlation between the rank orders for each model at significance level $\alpha << 0.01$.

The parametric model is recommended for use as a precursor to COBRA to narrow the number of contemplated closure installations to just those that require further, more detailed analysis and output.

A PARAMETRIC REGRESSION OF THE COST OF BASE REALIGNMENT ACTION (COBRA) MODEL

I. Introduction

Base Realignment and Closure

Following the end of the Cold War, the necessity of the United States' present level of military strength has come under scrutiny. This scrutiny is magnified by the onset of tightening budgets. Many United States officials including the Congressional Budget Office, Government Accounting Office, Office of Management and Budget, some Department of Defense officials, and numerous members of Congress are recommending to close military installations. "There appears to be a near consensus that a significant number of military bases should be closed as a cost savings and efficiency measure . . ."(Rasher, 1986:1).

Final responsibility for evaluating and selecting bases for closure falls to the Base Realignment and Closure (BRAC) Commission. Almand summarizes this process as follows:

> In an era of diminishing budgets, the Congress and the DoD are attempting to close unneeded military installations and realign remaining activities at others. The President's Commission on Base Closures and Realignment used mission and

> monetary criteria in their selection of possible closure candidates. (Almand, 1991:3)

BRAC employed a two-phase approach to incorporate mission-related and economic criteria for the identification and analysis of potential closure bases.

BRAC Process. Phase I accomplishes a complete inventory of installations, assigning them to categories such as operating ground troops, operating mobility aircraft, and operating submarines. A complete list of the twenty-two categories is given in Appendix A. Next, the military value of each installation is evaluated with respect to two criteria: (1) ability to complete assigned missions, and (2) capacity to accommodate additional missions (Defense Secretary, 1988;15-18). When an installation's mission is determined to be impaired according to these criteria, the Commission selects it for further detailed review during Phase II.

Phase II conducts an in-depth study to determine potential for relocation of activities or units. After development of relocation plans, a cost-estimating model is used to determine the costs and savings for the "best" relocation alternatives.

COBRA Model. Currently, the BRAC Commission uses the Cost of Base Realignment Action model (COBRA). This model, accepted by the General Accounting Office, provides an economic analysis of closure actions (COBRA v4.04, 1992:Program Documentation). For each closure scenario, execution of COBRA generates 20 detailed reports including appropriation details, mission costs, military construction costs, and personnel costs. The COBRA summary report condenses the detailed reports for use by decision makers. The summary report provides the payback period, 20-year net present value, and total one-time cost of the realignment action.

The payback period is

> . . . the point in time where savings generated equal (and then exceed) costs incurred. In other words, this is the point when the realignment/closure has paid for itself and net savings start to accrue. (COBRA v4.04, 1992:Program Documentation)

The 20 year net present value is

> ...the amount of dollars that would have to be invested during the Base Year at the assumed discount (interest) rate to cover the costs or match the savings at a specific point in the future. (COBRA v4.04, 1992:Program Documentation)

Net present value and payback period are the measures of effectiveness used to compare the economic consequences of one realignment scenario to another. COBRA requires thorough preparation and data collection by the user. In the simplest scenario, with transfer of assets and personnel from one closing base to a single gaining installation, the user may have to input 250 or more separately estimated values. Examples of input values include distances between the closing base and each gaining installation, tons of mission and support equipment to be transferred, costs of non-construction environmental mitigation, one-time unique costs associated with closure, and square footage of military construction at the gaining installation (COBRA v4.04, 1992:Program Documentation). Because of the many variables that require estimation, error can arise from a multitude of sources including faulty supplied data, lack of consistency, or incomplete information. This will generate a randomness in the COBRA Model itself simply due to entry data.

A Proposal for a Simplified Cost-Estimating Relationship

Using the COBRA model, the 1993 BRAC Commission evaluated 32 active duty Air Force bases, 14 Air Force Reserve installations, 10 Air National Guard installations, and 6 Air Force depots for possible closure. The decision to close a base is made in a sensitive political and socioeconomic environment, making it essential to work with reliable data. In addition to evaluating the possibility of closing Air Force bases, BRAC must also consider Army posts, Naval bases, Reserve installations, National Guard installations. Because of the enormity of the task and the aforementioned sources of input error, it would be desirable to build a simplified model that streamlines the BRAC process while maintaining a satisfactory level of accuracy.

A simplified model designed to capture the essential aspects of COBRA, requiring fewer input values and less data collection, could be developed to provide approximate estimates of base closure costs. The advantages of simplicity in the approximate model may more than outweigh the possibility of less accuracy relative to the detailed, but much more labor-intensive, COBRA model. If accuracy of the simplified model is judged acceptable, then it could be used in nearly all situations except when detailed cost justification by COBRA is necessary.

One type of simplified model that minimizes the number of inputs while maintaining accuracy is a cost estimating relationship (CER). A CER is a statistical model that treats cost to be a function of selected explanatory variables or parameters (Verma, 1987; 3). "Explanatory variables usually represent characteristics of system performance, physical features, effectiveness factors, or even other cost elements" (Fabrycky, 1991: 159).

Scope

This thesis develops a CER that estimates the costs of closing installations. To ensure compatibility of this model with COBRA, the CER must produce results that are consistent with COBRA. Chapter III describes the statistical procedures used to derive the CER and compare its accuracy to the full COBRA model. Linear regression lies at the heart of the procedure.

This thesis will be limited to modeling the 20 year net present value predicted by COBRA as a function of the selected parameters. Break-even year and payback period will not be modeled.

In contrast to COBRA, which generates economic measures of effectiveness using detailed data inputs, the proposed CER model will be a simplified model with a greatly reduced set of input variables. Decision makers will be able to use it to estimate, with minimum time and effort, the economic cost savings of closing a particular base.

Overview

To ensure equitable selection of bases for possible closure, the BRAC Commission requires an accurate and efficient means of forecasting base closure costs. The current method, which employs the COBRA model, is sound. Unfortunately, its numerous inputs necessitate that considerable data be collected and/or estimated. Because of the many inputs used in COBRA and the number of installations to evaluate, verification of data is a difficult task. Development of a parametric cost model as a discriminator for use early in Phase II deliberations will give valid results and will be less labor intensive. Its use can help narrow the field of potential closure bases providing time for verification of the data used in COBRA analysis.

Chapter II presents a review of the literature pertinent to base closure and the decision-making tools that are, or could be used, by BRAC. A synopsis of the BRAC Commission explains its process for selecting potential closure installations. Next is an exposition of the COBRA model, followed by a discussion of parametric models and their utility in economic cost modeling.

Chapter III describes the methodology employed in this thesis for development of a parametric cost estimating model. The basis for the techniques used is the least squares method of regression. Discussion of the least squares method in simple and multivariate regression leads to explanation of the model development procedure.

Chapter IV provides an analysis of the research results. The results of data gathering, screening, and preparation are discussed. Next, the results of the model generation and selection are given. Included are output of the stepwise and maximum R^2 techniques, and statistical verification of the selected best model. A comparison of the net present value predictions of COBRA and the best selected model concludes the chapter.

Chapter V concludes with a summary of work done and recommendations for future work.

II. Literature Review

This chapter presents a review of the literature that pertains to development of a parametric model of base closure costs. A description of the purpose and activities of the BRAC Commission, hereafter referred to as the Commission, is provided. A review of the Commission's work highlights the need for a simplified, yet reliable cost estimation tool. Because the Commission's work currently depends heavily upon the Cost of Base Realignment (COBRA) model, it receives consideration as the second section of this chapter. Next, a brief discussion of parametric models developed for other Air Force situations is presented. Finally, the chapter closes with a reiteration of the need for a parametric model to estimate base closure costs. The literature review reveals that such a model does not currently exist.

BRAC Commission

The Secretary of Defense chartered the Commission on Base Realignment and Closure on 3 May 1988 "to recommend military installations within the United States, its commonwealths, territories, and possessions for realignment and closure" (Defense Secretary, 1988:6). While the desire to reduce costs was an important reason for chartering the Commission, it decided that the "military value of each base should be the preeminent factor in making its decisions" (Defense Secretary, 1988:7). Thus, the selection of bases for closure is a two-phase process that takes into account both military value and cost reduction.

Phase I (Defense Secretary, 1988:15-16). The purpose of Phase I is to take an inventory of DoD installations and then screen them for detailed review during Phase II. The inventory of installations assures that none has been overlooked and, further, assigns each one to a functional category such as Operating Troops, Operating Aircraft, Headquarters, and Productions Facilities. All 22 functional categories assigned to each are given in Appendix A.

Following inventory, the installations are screened according to criteria that evaluate whether a base has the appropriate size to support current or future requirements and whether existing physical conditions are adequate for future purposes. Specifically, the Commission evaluates the military worth of an installation using 21 mission-related criteria (see Appendix B). Each criterion measures an attribute of the base, assigning one of three ratings: marginal, acceptable, and fully satisfactory. The criteria for ratings are well-defined and consistently applied.

Following assignment of the 21 ratings, they are weighted and summed to yield and overall rating. From the rank-ordered list that results, the Commission selects the lowest-ranking bases for detailed review in Phase II.

Phase II (Defense Secretary, 1988:16-18). During Phase II, the Commission develops options for closure and realignment of the lowest-ranking installations identified in Phase I. Relocation of mission-essential units requires identification of appropriate gaining installations. In this process, the first step is to identify all activities that need to be moved and then to develop relocation options. The preferred relocation option for each activity is the one that maximizes mission attainment. The Commission then employs the COBRA cost-estimating model to evaluate all potential relocation options. COBRA output provides a detailed analysis of the costs and savings associated with each relocation option. Output measures of merit include net present value and payback period.

The military services participate actively during the analyses. They provide data pertaining to the physical characteristics of installations. Each service also contributes expertise regarding current missions, the relevance of proposed criteria by which to evaluate existing missions, and the weights to be assigned to each of the mission-related criteria. In addition, the services help to identify the activities to be relocated and propose to the Commission which installations should gain the relocated activities.

COBRA Model

COBRA is a cost model that utilizes basic principles of economic analysis. Dover and Oswald define economic analysis as

> (a) a systematic approach to the problem of choosing how to employ scarce resources and (b) an investigation of the full implication of achieving a given objective in the most efficient and effective manner. (Dover and Oswald, 1974:47)

COBRA models closures, deactivations, and realignments of DoD installations (Brown, 1989:3):

- Closure - All activities are transferred away from the losing base and the property is sold. Some costs are incurred to prepare the base for sale.
- Deactivation - Most of the activities are transferred away from the losing base, and a caretaker force is left in place to provide a minimal maintenance and security capability.
- Realignment - Some activities are transferred away from the losing installation but it continues to operate.

COBRA uses standard cost factors to convert the financial impact of base closures, deactivations, or realignments into dollar values (inflated relative to the base year). Cost factors may differ between the services and may also reflect local conditions that are a function of geographic area. The output of the model is expressed in terms as payback period and 20-year net present value.

The costs and savings considered for data input are categorized as either one-time or recurring. One-time costs and savings include construction, administrative planning and support, personnel retirements and severances, personnel relocation costs, equipment freight and transport costs, environmental mitigation, land purchases and sales, and shutdown costs. Recurring costs and savings include housing allowances, increased Civilian Health and Medical Program of the Uniformed Services (CHAMPUS) costs, caretaker costs, salary savings, and changes in base overhead and maintenance costs (Nelson, 1989:39,40 and Defense Secretary, 1988:51). Table 2.1 lists examples of estimated data inputs required to use COBRA.

In the FY 1993 COBRA analysis of bases, 32 Air Force bases were examined for closure. To complete the realignment scenario for each base, gaining installations (of personnel and equipment) had to be designated. Each scenario averaged four gaining installations, which requires approximately 700 inputs.

A major drawback of the COBRA model is the number of inputs required. To use the model correctly in the case of a single gaining installation when a base closes, the user may have to input as many as 300 values. Estimates are required for all of the inputs.

TABLE 2.1

ILLUSTRATIVE ESTIMATES REQUIRED BY THE COBRA MODEL (Nelson, 1989:24,25)

EQUIPMENT COSTS

- Amount of material required at gaining base
- Condition of the equipment and ability to be moved
- Size, weight, and quantity
- Mode of transportation
- Shipping distance
- Amount of excess material to be routed to local bases
- Amount of unusable equipment to be turned in to salvage

CIVILIAN WORKFORCE REDUCTION
(reimbursable expenses)

- Per diem
- Mileage
- Travel for one round trip house hunting excursion
- Temporary quarters subsistence
- Broker's fees, real estate commissions, and miscellaneous expenses for home sale and repurchase
- Transportation and storage of household goods
- Relocation income tax

Cost Estimating Models

Within the Department of Defense several generally accepted economic analysis techniques exist. Among them are detailed estimating, cost estimating relationships (CER), expert opinion, and analogy (Verma, 1987:2). None can be determined to be best; their appropriateness is situational.

Detailed estimates identify activities at a very detailed level and sum up the costs to higher levels of aggregation. CERs, sometimes referred to as parametric models, use statistical methods which treat cost as a function of selected variables called cost drivers (Verma, 1987; 3). In the expert opinion approach, the cost analyst seeks the opinion of experts about the system being analyzed. In the method of analogy, comparisons are made of the proposed system similar existing systems. Costs are adjusted according to differences in the systems (Verma, 1987:2-3).

A review of the literature (Brown, 1989; Nelson, 1989; Gatlin, 1992; Holk, 1989; Olver, 1991; Rasher,1986) indicates that most base closure models are of either the detailed analysis or accounting type. These models estimate cost of closure by disaggregating and describing the closure process itself. They normally include a large number of variables corresponding to the many activities that make up the closure effort. Advantages of these models are that (1) costs can be tracked more easily to provide better management of dollars, and (2) the amount of detail lends itself to sensitivity analysis and tradeoff decisions. Disadvantages are that (1) the models require detailed information, and (2) they utilize many variables that are estimated (Verma, 1987:4).

Because of the sensitivity of closure actions, detailed reports are helpful in explaining decisions. The "black box" nature of a parametric estimating function does not necessarily provide a logical explanation of the final output. Nonetheless, it can be a powerful tool if used properly:

> If a model could estimate costs equally well while using significantly less variables, the purpose would be served just as well. Cost estimating relationships can do just that. (Verma, 1987:5)

Verma [1987] compared an accounting model to a parametric model in the same scenario. The accounting model had fifty-one variables while the parametric model had only seven with at least the same amount of predictive accuracy.

Cost estimating relationships are developed by means of regression analysis. Regression analysis, employing the least squares method, is a statistically proven method which establishes a functional relationship between variables in order to predict the value of the dependent variable (cost) on the bases of the independent variables (cost drivers). The functional relationship is a linear equation (Cain, 1993)

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_K X_K + \varepsilon .$$

Y is the cost, β_i is the i-th parameter coefficient (estimated by *regression* analysis), X_i is the *i*-th parameter, *K* is the number of parameters, and ε is the error term.

There are two myths about the degree of accuracy and level of detail necessary for a useful model.

MYTHS (Fisher, 1971:76):

- We must strive for a high degree of accuracy in an absolute sense.

- A higher degree of accuracy can be attained by going into a greater amount of detail.

Verma exposed the first myth as follows (Verma, 1987:6-7):

> In the context of long range planning, the possibility of accomplishing a high degree of accuracy in the absolute sense is remote. This is so because of the characteristics of long range planning. These characteristics include uncertainties, <u>lack of detailed information and data</u>, and a wide range of alternatives. Under these conditions, highly accurate cost estimates are most unlikely. This is not critical because long term planning efforts require relative comparisons between alternatives. <u>If the cost estimates provide sufficient information to facilitate the best decision, then they have served their purpose</u>. Analytical cost estimating techniques which treat alternatives consistently are better suited for comparative cost analyses. [emphasis added]

Verma brings to light the second myth, as well (Verma, 1987:7)

> [It] is generally not true and particularly false in the context of long range planning. Under conditions of knowledge gaps and paucity of data, to force the analysis into a finer and finer grain of detail will force the analyst into essentially using fictitious numbers to fill in the categories that are overly detailed.

Parametric models have been developed for several situations in the DoD. One of the most reliable metal airframe parametric cost models is the Rand Corporation *Development and Procurement Cost of Aircraft* model (DAPCA III) (Kage, 1983:6-7). An historical database of performance characteristics costs for aircraft was used to develop the model. DAPCA III uses two performance parameters, Aeronautical Manufacture's Planning Report (AMPR) weight and maximum speed in knots at best altitude, to determine the cost of future airframes.

Other parametric airframe cost models are the *Planning and Research Corporation* (PRC) model, the *Science Applications Incorporated* (SAI) model, the *J. Watson Noah* (JWN) model, and the Air Force Flight Dynamics Laboratory *Modular Life Cycle Cost Model* (MLCCM) (Kage, 1983:22). Table 2.2 highlights the characteristics of each model.

TABLE 2.2

AIRFRAME COST
MODEL PARAMETERS (Kage, 1983:25-31)

MODEL	PARAMETERS
PRC	responsible agency, lot quantity, delivery rate, airframe weight growth, speed at best altitude, speed at sea level, altitude, AMPR weight, aircraft empty weight
SAI	weights by 1) section -- wing, tail, etc. and 2) subsystems -- avionics, electrical systems, air conditioning, etc.
JWN	maximum speed at best altitude, AMPR weight, gross takeoff weight/AMPR weight (ratio), a dummy variable to account for complexity
MLCCM	wing thickness/chord (ratio), wing area, fuselage wetted area, speed, ultimate load limit, max mach speed, max gross weight, total wetted area, number of prototype aircraft, etc.

Further examples of parametric models developed in the Air Force include Dupre's model for predicting direct costs or cost changes for electronic systems flight tests (Dupre', 1983:11), Gardner and Passarello's model to predict first unit costs of flight simulators by using certain common characteristics (Gardner, 1981:12), and Muenchow's model for estimating cost of buying out telecommunications switches that are on leasing agreements (Muenchow, 1991:ii).

TABLE 2.3

AIR FORCE DEVELOPED PARAMETRIC MODELS
(Dupre', 1983:47; Gardner & Passarello, 1981:37; Muenchow, 1991:34,41,46)

RESEARCHER	PARAMETERS
Dupre'	number of sorties, number of months, three dummy variables for type of aircraft, two dummy variables for type of test
Gardner & Passarello	cooling capacity, weight, degrees of freedom, rate of power consumption, computer instruction processing speed
Muenchow	LARGE SWITCHES: monthly lease fee, age SMALL SWITCHES: line count, monthly lease fee, age

III. Methodology

COBRA completes an economic analysis of base closure actions. Because of the many variables to be estimated, the data required by COBRA may be prone to errors and inconsistencies. In contrast, parametric models minimize the number of variables yet provide a great deal of accuracy. This thesis develops a parametric cost estimating model that can be used to predict closure costs without resorting to the enormous data collection effort required by COBRA.

This chapter presents the methodology to be used for the development of a parametric base closure cost model. First is a brief discussion of simple (i.e., one variable) linear regression using least squares estimation. Application of linear regression involves the principal assumption that the unknown relationship between COBRA inputs and outputs can be fitted to a line. "The unknown parameters [of the line] are estimated under certain other assumptions with the help of available data, and a fitted equation is obtained" (Draper & Smith, 1981:2).

The case of one-variable regression leads to a derivation of the multivariate or matrix approach to linear regression. Many statistics and regression analysis texts (Devore; Draper & Smith; Gunst & Mason; Milton & Arnold;

Montgomery & Peck; Sen & Srivastava) dedicate entire chapters each to simple and matrix regression analysis.

Following the discussion of theory, the remainder of the chapter addresses the implementation of linear regression in the context of model development. An important point to keep in mind throughout the model building process is that absolute truth, or perfect prediction, cannot be achieved. The values used in COBRA are estimates themselves; therefore, COBRA's predictions are only as good as the estimates.

The objective of the least squares method is to minimize the sum of squared errors in the predictions. The goal of the model development techniques employed in this thesis is to minimize the number of variables and provide accurate predictive capability -- explain the most with the least. A model will be developed which will make a prediction using as few predictor variables as possible while minimizing the sum of squared errors.

The first step in model development is to gather the data, analyze by inspection, and complete some preparation of the data. The inspection process can narrow the data field to those input variables thought to be most significant to the response variable. Once the field of input variables is narrowed, the correlations between all remaining input variables are calculated. Analysis of the correlation coefficients provides the basis for variable preparation and elimination.

Once the most significant predictor variables are determined, model selection techniques are applied. The techniques discussed in this chapter are stepwise selection and maximum R^2. Proper application of these techniques will provide a method for determining the optimum number of variables and "best" model. The t-test is also discussed as a method for checking the significance of individual regressors in a model.

Once the best model, or models, are selected, they are verified by examining residuals and checking multicollinearity diagnostics.

Simple Linear Regression.

The estimation procedure to be used here is based on simple linear regression and the least squares method. "The simplest deterministic mathematical relationship between variables x and y is a linear relationship $y=\beta_0+\beta_1 x$ " (Devore, 1991:454). In this equation, β_1 is the slope of the line and β_0 is the y-intercept. Customarily x is called the independent variable and y is called the dependent variable. To avoid confusion with statistical independence, x is sometimes referred to as the *predictor* or *regressor* variable and y as the response variable. Both sets of terminology will be used in this thesis.

When a straight line is fitted to the data, the underlying assumption is that the data follows a linear pattern of the form $y=\beta_0+\beta_1 x$. However, the data rarely fits this pattern exactly. Thus the relationship becomes $y_i=\beta_0+\beta_1 x_i+\varepsilon_i$, where $i=1,2,...,n$ (n is number of data points) and ε_i is referred to as the random error term. "Without ε, any observed pair (x,y) would correspond to a point falling exactly on the line $y=\beta_0+\beta_1 x$, called the true regression line" (Devore, 1991: 456). The random error term is the difference between the observed value of y_i and the straight line $\beta_0+\beta_1 x_i$. "It is convenient to think of ε as a statistical error; that is, it is a device that accounts for the failure of the model to fit the data exactly" (Montgomery & Peck, 1982:2-3).

In conducting a regression study, paired sets of data (y_i, x_i), or observations, are used to estimate the parameters β_1 and β_0. It is important to note that the true values for β_1 and β_0 will never be known. Because of this, the same is true for ε_i. These values will be approximated from the available data. Letting b_0, b_1, and e_i denote the estimates for β_1, β_0, and ε_i respectively, the estimated line of regression takes the form

$$y_i=b_0+b_1 x_i+e_i \qquad (i=1,\dots,n) \tag{1}$$

The term e_i is called the *residual*. Figure 3.1 graphically illustrates the difference between ε_i and e_i.

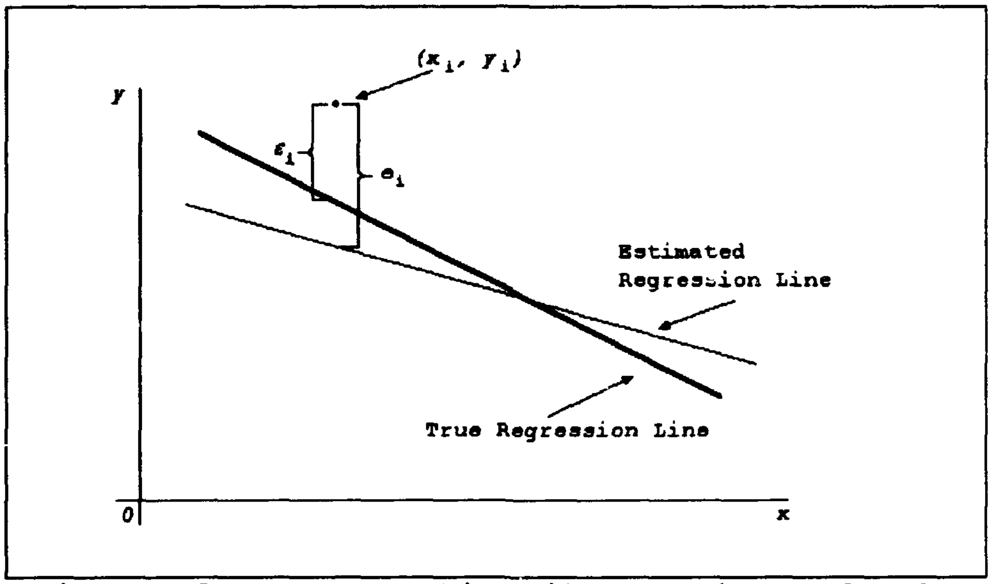


Figure 3.1. Graphic illustration of the difference between the residual, e_i, and the random error term, ε_i. (Milton & Arnold, 1986:340)

Least Squares Estimation

The parameters β_1 and β_0 are estimated by the method of least squares. This method employs the idea that from the many lines that can be drawn through a plot of the observations (x_i, y_i), the one that "best fits" the data can be selected. One common measure of best fit occurs when the values of b_0 and b_1 minimize the sum of the squares of the residuals. "In this way, we are essentially picking the line that comes as close as it can to all data points simultaneously" (Milton & Arnold, 1986:340).

The residual e_i is sometimes called the "residual error." The sum of the squares of the residuals is often then called the "error sum of squares" or Sum of Squares

Error (SSE). The following derivation of the least squares estimates for β_1 and β_0 is adapted from Milton & Arnold, 1986:341. By manipulation of equation 1, SSE is given by

$$SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2 \tag{2}$$

Differentiate SSE with respect to b_0 and b_1 to obtain

$$\frac{\partial SSE}{\partial b_0} = -2\sum_{i=1}^{n} (y_i - b_0 - b_1 x_i) \tag{3}$$

$$\frac{\partial SSE}{\partial b_1} = -2\sum_{i=1}^{n} (y_i - b_0 - b_1 x_i) x_i \tag{4}$$

Set both partial derivatives equal to zero and use the rules of summation to obtain the equations

$$n b_0 + b_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \tag{5}$$

$$b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \tag{6}$$

These equations are called the *normal equations*. They can be solved to obtain the following estimates of β_1 and β_0 (Devore, 1991:461).

$$b_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{7}$$

$$b_0 = \frac{\sum_{i=1}^{n} y_i - b_1 \sum_{i=1}^{n} x_i}{n} = \bar{y} - b_1 \bar{x} \tag{8}$$

<u>Gauss-Markov Conditions</u> (Sen & Srivastava, 1990:11-13; Draper & Smith, 1981: 22-23)

Thus far the least squares method has taken for granted some important assumptions or conditions. If these conditions, called Gauss-Markov conditions, are met, the least squares method gives good predictions for β_1 and β_0. The conditions are as follows:

1. ε_i is a random variable with mean zero and variance σ^2 (unknown), that is, $E(\varepsilon_i)=0$, $V(\varepsilon_i)=\sigma^2$.

2. ε_i and ε_j are uncorrelated, $i \neq j$, so that $\mathrm{cov}(\varepsilon_i,\varepsilon_j)=0$ or $E(\varepsilon_i\varepsilon_j)=0$.

3. ε_i is a normally distributed random variable, with mean zero and variance σ^2 by (1), that is, $\varepsilon_i \sim N(0,\sigma^2)$. Under this assumption, ε_i, ε_j are not only uncorrelated but necessarily independent.

When these conditions are met, they assure that a "best linear unbiased" estimator occurs which is also a "maximum likelihood" estimator (Cain, 1993).

Coefficient of Determination

The error sum of squares, SSE, can be interpreted as a measure of how much variation in *y* is left unexplained by the selected regression equation or model.

$$SSE = \sum (y_i - \hat{y}_i)^2 = \sum \left[y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 x_i \right) \right]^2 \tag{9}$$

A quantitative measure of the total amount of variation, both that which is explained and not explained by the model, in observed *y* values is given by the total sum of squares

$$\mathrm{SST} = \sum_{i=1}^{n} (y_i - \overline{y})^2 = \sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i \right)^2 \Big/ n \tag{10}$$

The ratio SSE/SST is the proportion of total variation that cannot be explained by the simple linear regression model, and 1 - SSE/SST is the proportion of observed *y* variation explained by the model. The coefficient of determination, R^2, is given by

$$R^2 = 1 - \frac{SSE}{SST} \tag{11}$$

The higher the value of R^2, the more successful the model is in explaining the variation of *y*. The statistic R^2 should be used with caution, since it is always possible to make R^2 large by simply adding terms to the model.

> Although R^2 will always increase as more variables are added, the mean squared error usually will first decrease and then increase as additional variables are selected. Typically, the experimenter selects the model corresponding to the smallest mean squared error. (Milton & Arnold, 1986:428)

The mean squared error is defined by dividing the error sum of squares by its degrees of freedom, $n-(k+1)$ (n is the number of observations and k is the number of variables in the model).

Analysis of Variance

Analysis of Variance (ANOVA) is a statistical methodology in which the total variation in a measured response is partitioned into components which can be attributed to recognizable sources of variation. The components are usually summarized in an ANOVA table (Table 3.1).

The total sum of squares (SST) and error sum of squares (SSE) were defined earlier. By understanding that SSE accounts only part of the total variance SST, the remaining variance can be accounted for by the equation SST - SSE = remaining variance. This remaining variance is the variance due to the regression and is naturally termed the regression sum of squares or SSR.

Additional components given in an ANOVA table are the mean squares, which are simply the sums of squares divided by their respective degrees of freedom (df). The df for the regression is the number of variables in the model, k. The regression mean square is given by: MSR = SSR/k. The df for the error is $n-(k+1)$. The error mean square is given by: MSE = SSE/(n-k-1).

The functions MSR and MSE (also designated by MS_{Reg} and s^2, respectively) have their own distribution, mean, and variance. "Also, since these two random variables are independent, a statistical theorem tells us that the mean-square ratio

$$F = \frac{MS_{Reg}}{s^2}$$

follows an F distribution...." (Draper & Smith, 1981:32). This statistic, given in the ANOVA table, can be used in multivariate analysis (which will be explained later) to test the overall regression equation. The test is actually to determine whether all the parameters ($\beta_1 \ldots \beta_k$) are statistically equal to zero (the null hypothesis H_0). With a specified risk level, α, a selected percentage point of the F-distribution can be compared against the calculated mean-square ratio.

If the mean-square ratio exceeds the critical F-value, $F(k,n-k-1,\alpha)$, a "statistically significant" model has been found. This does not necessarily mean that the model is the best, or even a good predictive model. It just tells the experimenter that the

> . . . proportion of the variation observed in the data, which has been accounted for in the equation, is greater than would be expected by chance in $100(1-\alpha)\%$ similar sets of data with the same values of n and $\mathbf{X}$. (Draper & Smith, 1981:93)

Work by J.M. Wetz suggests that in order for a model to be a good predictor, the mean-square ratio should exceed the selected percentage point from the F-distribution by at least four times (Draper & Smith, 1981:93). For example, if $k = 9$, $n = 50$, $\alpha = 0.05$, $F(9,40,0.95) = 2.12$. The observed mean-square ratio (F-ratio) would have to exceed 8.48 for the model to be a satisfactory prediction tool. Table 3.1 below gives a symbolic representation of the variance components provided in an ANOVA table.

TABLE 3.1

SYMBOLIC ANALYSIS OF VARIANCE (ANOVA) TABLE
(Gunst & Mason, 1980:157)

SOURCE	Deg. of Freedom	Sum of Squares	Mean Squares	F-Ratio	R^2
Regression	k	SSR	MSR	MSR/MSE	SSR/SST
Error	n-k-1	SSE	MSE		
Total	n-1	SST			

Multivariate Regression

Regression analyses utilizing several predictor variables are simultaneously more flexible and more complicated than those using a single predictor variable. Flexibility comes from the researcher's opportunity to assess the relationship between the dependent variable and several independent variables, several specifications of one independent variable, or a combination of both (Gunst & Mason, 1980:128). In contrast, in simple linear regression, a single predictor variable is forced to adequately fit the response variable.

The tradeoff of using multiple-variable regression is complexity. Most multivariate analysis requires the use of a computer. If statistical analysis packages such as SAS (Statistical Analysis System: SAS Institute, Inc. Raleigh, NC), SPSS (Statistical Package for the Social Sciences, McGraw-Hill), BMD (Biomedical Computer Programs, University of California Press), or MINITAB (Deerbury Press) are available, many analysis techniques can be applied very simply (assuming some knowledge of the software). However, selecting the "best" model requires much more than running computer programs. The complexity of multivariate analysis comes in interpreting the output. The researcher must not only "ask" the computer for the appropriate output, but must also correctly discern the implications of the output.

Though there are many rules to aid the analyst, determining the optimum prediction equation often relies on experience and judgment. "Selecting a final model is, in many ways, an art rather than a science" (Milton & Arnold, 1986:424).

The basis for the statistical techniques used in multivariate regression is also the least squares method. Matrix algebra can be used to concisely represent the least squares estimators for multiple-variable prediction equations. The multiple linear regression model $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + \varepsilon_i$ (i = 1, 2, ..., *n*) can be written as

$$\bar{\mathrm{Y}} = \mathbf{X}\bar{\beta} + \bar{\varepsilon} \tag{12}$$

where,

$$\bar{\mathrm{Y}} = \begin{bmatrix} \mathrm{Y}_1 \\ \mathrm{Y}_2 \\ \vdots \\ \mathrm{Y}_n \end{bmatrix} \qquad \bar{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \qquad \bar{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & x_{31} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & x_{32} & \cdots & x_{k2} \\ 1 & x_{13} & x_{23} & x_{33} & \cdots & x_{k3} \\ \vdots & & & & & \\ 1 & x_{1n} & x_{2n} & x_{3n} & \cdots & x_{kn} \end{bmatrix}$$

Note that $\bar{\mathrm{Y}}$ is the response vector, $\bar{\beta}$ is the parameter vector, and $\bar{\varepsilon}$ is the random error vector. $\mathbf{X}$ is referred to as the *model specification matrix* (Milton & Arnold, 1986:402). The dimension of $\mathbf{X}$ is *n* x (*k*+1). The column of *n* ones accounts for the intercept term, and the other cells represent the observations of the predictor variables for the model.

As in simple linear regression, the intention of using the least squares method is to minimize the sum of the squared residuals. The residual vector $\bar{r}$ is defined to be

$$\bar{r} = \bar{Y} - \hat{\bar{Y}} = \bar{Y} - \mathbf{X}\hat{\bar{\beta}}$$

so then the sum of the squared residuals is expressible as (Gunst & Mason, 1980:133)

$$\sum_{i=1}^{n} r_i^2 = \bar{r}'\bar{r} = \left(\bar{Y} - \mathbf{X}\hat{\bar{\beta}}\right)'\left(\bar{Y} - \mathbf{X}\hat{\bar{\beta}}\right) \qquad (13)$$

Minimization of equation 12 with respect to $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ yields the least squares estimator (Gunst & Mason, 1980:133)

$$\hat{\bar{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\bar{Y}. \qquad (14)$$

The resulting vector gives the predicted values for the parameters of the model.

The matrix operations involved in multivariate least squares regression can be cumbersome. Fortunately, the software packages mentioned earlier will effortlessly do the work. The SAS package will be used in this thesis, although other commercial software could have been used equally as well.

Data and Regression Analysis

The model development process outlined below is certainly not the only way to build a parametric model. The techniques and methods used follow "hard and fast" rules; however, the "human element" is important. The experience, intuition, and knowledge of the researcher is critical in every phase of the process. There is no perfect procedure. The "best" steps to be taken in model development often depend on the goal of the regression analysis and the judgment and preferences of the researcher. The process used in this thesis is a collaboration of the research effort of the authors and the experience of local experts.

Data Inspection

Prior to application of statistical techniques, the researcher should attempt to get a good "feel" for the data. Understanding where the data comes from and how it will be used may make application of statistical techniques and model selection procedures much easier later in the analysis. Data inspection may reveal norms and inconsistencies which may be important later on. In development of a cost model, the variables most significant to the output, or cost drivers, should be identified. Examination of the data can be the first step to identifying some of the most and least significant cost drivers. Simple techniques that can be used to prepare data include forming aggregates of various variables and eliminating others. This step need not be extensive and time-consuming. The researcher will continue to learn more about the data as the analysis continues and may have to return to this step.

Correlation

The model selection procedures described below generally provide good models. However, the capability of the models produced often depends on the quality of the data used. One characteristic of some data sets that may cause problems is high correlation among the independent variables (Sen & Srivastava, 1990;240-242). This problem can be eliminated or reduced by the researcher early in the data examination phase of the process by conducting correlational study to further distill the variable field.

Correlation measures the closeness of a linear relationship between two variables. The value used to examine the correlation between two variables is known as the Pearson correlation coefficient (Milton & Arnold, 1986:364). It is defined by

$$\rho_{x_i x_j} = \frac{\operatorname{cov}(X_i, X_j)}{\sqrt{\left(\operatorname{var}(X_i)\operatorname{var}(X_j)\right)}}$$

This represents the true product-moment correlation between two variables. The sample correlation estimates the true correlation. It is given by (SAS, 1985:862)

$$r_{x_i x_j} = \frac{\sum_{q=1}^{n}(x_i - \bar{x}_{iq})(x_j - \bar{x}_{jq})}{\sqrt{\left(\sum_{q=1}^{n}(x_i - \bar{x}_{iq})^2\right)\left(\sum_{q=1}^{n}(x_j - \bar{x}_{jq})^2\right)}} \tag{15}$$

If two variables are highly correlated, one or both can be manipulated (possibly eliminated) before further analysis continues. The SAS CORR procedure will produce the correlation coefficients between variables.

Model Selection

The goal in this thesis is to develop a cost model that will be as easy to use as possible. Fulfillment of this goal will entail (1) selecting variables that minimize the estimation effort on the part of the user and (2) minimizing the number of variables in the model. The variables, and aggregate variables, chosen to be significant cost drivers may or may not require extensive estimation. The decision to include them in further analysis is left to the judgment of the researcher. The latter objective can be achieved through the application of a number of techniques. There is no single statistical procedure that will absolutely select the "best" regression equation. Best in this case is defined by most predictive capability with least number of variables. The two techniques described below, stepwise selection and maximum R^2, can aid the researcher in choosing the optimum number of variables and eventually the best model. These techniques were chosen for this thesis (rather than others such as forward selection, backward elimination, and use of the R^2 statistic) because they appear to complete a more rigorous check of the combinations of different variables. However, they will not necessarily produce the same best model. Once again, selection of a technique is left to the judgment of the researcher. In any case, model selection should be followed by model verification, which is discussed later. Both procedures discussed below can be carried out by SAS.

Stepwise Selection. The stepwise (SW) selection method begins with the basic model $Y=\bar{Y}$. The predictor variable having the most correlation with Y, highest partial correlation coefficient, is then added to the model. A regression of Y on the new variable is completed and the variable is checked for significance using the F-test. If the variable is statistically significant, another variable is added. Selection of new variables is also based on the partial correlation coefficients; however, the coefficients must be recalculated each time a variable is added to the model to account for the "loss" of the variable from the selection set. Y is regressed on the new set of variables.

The new model is then checked for significance: improvement in R^2 is noted and the partial F-values for each variable now in the equation is examined. The lowest partial F-value is compared with an appropriate F-percentage (F exit). "This provides a judgment on the contribution of the least valuable variable in the regression at that stage . . ." (Draper & Smith, 1981;308). The corresponding predictor variable is either retained in the model or rejected according to the F-value comparison. If rejected, the variable is returned to the selection set for possible inclusion at a later stage.

Once again, a new predictor variable is chosen based on its adjusted partial correlation coefficient. This new variable is checked to see if its partial F-value exceeds a preselected F-percentage for entry into the model (F entry). Following a new regression of *Y*, a test of the F criterion for each variable is completed again. "Each time a new variable is entered into the model, all the variables in the previous model are checked for continued importance" (Milton & Anderson, 1986:427). Basically, the least significant predictor is tested at each stage of the stepwise procedure. Once no variables can be removed from the equation and the next best candidate variable cannot hold its place in the equation, the process stops.

Maximum R^2. This procedure attempts to find the best *k*-variable equation, starting with a *k-1* variable equation. The regressor providing the greatest increase in R^2 is added. Given the new equation, a check is made to see if switching a regressor currently in the model with one currently excluded will increase the value of R^2. If so, the switch is made and the process continues until value of R^2 can no longer be increased (Montgomery & Peck, 1982:279). The advantage to using the maximum R^2 method is that it allows the researcher to examine the summary statistics for a larger number of models than would ordinarily be generated by the stepwise method.

t-test

The t-test is not a model selection technique, but this statistic can be helpful when selecting model parameters. The t-statistic allows the researcher to test the significance of a regression coefficient in the model. The hypotheses for testing any regression coefficient are $H_0: \beta_i = 0$ and $H_0: \beta_i \neq 0$. If $H_0: \beta_i = 0$ is not rejected, then the regressor x_i is considered statistically equally to zero and can be eliminated (Montgomery & Peck, 1982:132). The test statistic for this hypothesis is

$$t_0 = \frac{\hat{\beta}_i}{\sqrt{\hat{\sigma}^2 C_{ii}}} \tag{15}$$

where C_{ii} is the diagonal element of $(\mathbf{X}'\mathbf{X})^{-1}$ corresponding to $\hat{\beta}_i$. The null hypothesis $H_0: \beta_i = 0$ is rejected if $|t_0| > t_{\alpha/2, n-k-1}$. Note this is a test of the contribution of x_i given the other regressors in the model.

Model Adequacy

The assumptions that have been made thus far in the model analysis are:

1. The relationship between y and x is linear or well-approximated by a straight line.

2. The error term ε has zero mean.

3. The error term ε has constant variance σ^2.

4. The errors are uncorrelated.

5. The errors are normally distributed.

Gross violations of these assumptions may yield an unstable model. Following are a few final methods for checking the adequacy of the model.

Residual Analysis. Assumption number 3 above summarizes one of the Gauss-Markov conditions which states $V(\varepsilon_i) = V(y_i)$ is a constant, σ^2. Violation of this condition is often called heteroscedasticity (Sen & Srivastava, 1990; 111). Heteroscedasticity exists when the error terms, or residuals, have unequal variances. One method of detecting heteroscedasticity is to examine the residuals for systematic patterns (Gunst & Mason, 1980;237).

Residuals are defined as $e_i = y_i - \hat{y}_i$, $i = 1, 2, \ldots, n$. The residual may be viewed as a measure of variability not explained by the regression model.

> Examining the residuals is one of the most important tasks in any regression analysis. A residual analysis involves the careful inspection of the differences between the observed and predicted values of the response variable after a prediction equation is fit to the data. In doing so, one hopes to spot any anomalies in the data which might cause poor prediction or poor parameter estimation. (Gunst & Mason, 1980:220)

After viewing the residuals, the researcher should be able to conclude (1) the assumptions mentioned above appear to be violated, or (2) the assumptions do not appear to be violated. Conclusion number two does not necessarily mean that the assumptions are all correct; "it means merely that on the basis of the data we have seen, we have no reason to say that they are incorrect" (Draper & Smith, 1981:142) The objective in plotting residuals is chiefly for observing patterns in the data.

The following are graphical methods for examining the residuals. With the use of good statistical packages they are easy to do and "are usually very revealing when the assumptions are violated" (Draper & Smith, 1981:142). The principal residual plots are

- Frequency distribution of the overall residuals
- The residuals (or Studentized residuals) against the fitted values Y_i
- The residuals against the independent variables

An overall plot can be done in the form of a histogram when the number of residuals is high. If the model is correct, the histogram should have some resemblance to a normal curve. It is often beneficial to plot the Studentized (standardized) residuals rather than the error terms themselves (Draper & Smith, 1981;144: Sen & Srivastava, 1990; 102). The Studentized residual is simply the residual divided by the standard error.

The plots of the residuals against the fitted values and against the independent variables should represent a "horizontal band." Abnormality in the residuals (heteroscedasticity) may cause the plots to take on the shape of (1) an expanding band, (2) an inclined band, or (3) a curved band. Examples are shown in Figure 3.2 below. These plots would indicate respectively the need for extra terms, error in the analysis, and the variance is not constant as assumed (Draper & Smith, 1981:147-148).

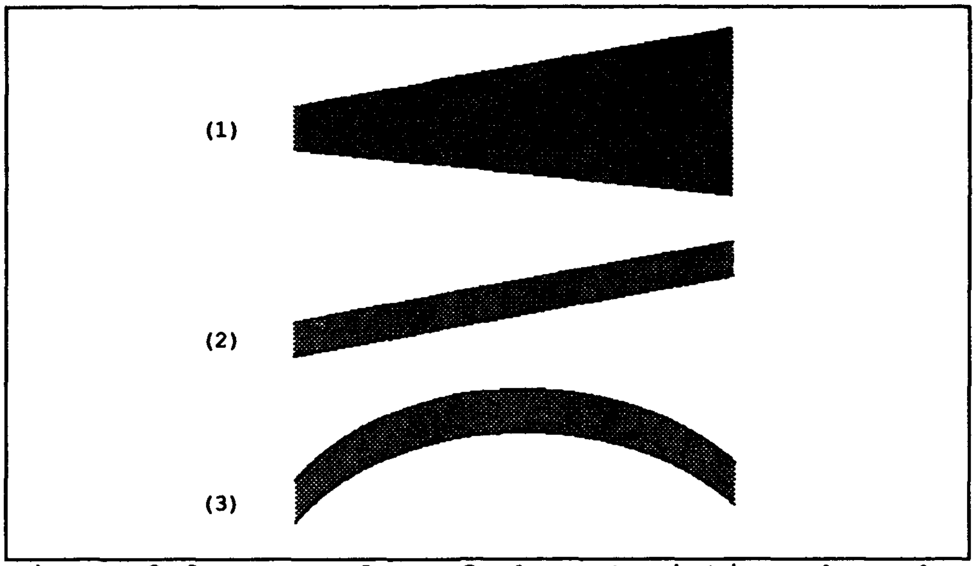


Figure 3.2. Examples of characteristics shown by unsatisfactory residuals behavior. (Draper & Smith, 1981;146)

Multicollinearity. The quality of estimates, as measured by their variances, can be seriously and adversely affected if the independent variables are closely related to each other. When there are near linear dependencies between regressors, the problem of multicollinearity is said to exist. Two main sources of multicollinearity are (1) more parameters are postulated than are needed to express the data, or (2) the data are not adequate to estimate the model (Draper & Smith, 1981:258).

Multicollinearity can be checked through use of several diagnostics. The one used in this thesis is the variance inflation factor method. Recall that $\mathbf{C} = (\mathbf{X}'\mathbf{X})^{-1}$. The *i*-th diagonal element of $\mathbf{C}$ can be written as $C_{ii} = \left(1 - R_i^2\right)^{-1}$, where R_i^2 is the coefficient of determination obtained when x_i is regressed on the remaining k-1 regressors. $C_{ii} = \left(1 - R_i^2\right)^{-1}$ has been termed the variance inflation factor VIF (Montgomery & Peck, 1982:300).

> The VIF for each term in the model measures the combined effect of the dependencies among the regressors on the variance of that term. One or more large VIFs indicate multicollinearity. Practical experience indicates that if any of the VIFs exceeds 5 or 10, it is an indication that the associated regression coefficients are poorly estimated because of multicollinearity. (Montgomery & Peck, 1982:300)

Model Adequacy vs. Model Validation. The adequacy of the model is checked by methods such as examination of the residuals and checks for multicollinearity. The adequacy shows the fit of the regression model to the available data. Model validation, on the other hand, is more directed to determining if a model will function properly for its intended use. The fact that a model fits the data used in its development does not necessarily mean that it will be successful in its final application. Several differences could exist in the data used during regression and that used in the field. Also, the developer has little control over use of the model once it is given to the user.

Listed below are three procedures useful in validating regression equations (Montgomery & Peck, 1982;425-426).

1. Analysis of the model coefficients and predicted values including comparisons with prior experience, physical theory, and other analytical models, or simulation results.

2. Collection of fresh data with which to investigate the model's predictive performance.

3. Data splitting; that is, setting aside some of the original data and using these observations to investigate the model's predictive performance.

Summary

The least squares method, employed in both simple and multivariate regression, minimizes the sum of the squares of the residuals. Assuming the following conditions are met, the least squares method will produce a good linear fit of the linear regression to the data.

1. The relationship between y and x is linear or well-approximated by a straight line.

2. The error term ε has zero mean.

3. The error term ε has constant variance σ^2.

4. The errors are uncorrelated.

5. The errors are normally distributed.

Once the data is examined and prepared by judgment on the part of the researcher, several variable selection techniques can be applied to aid in developing the "best" model. This thesis will narrow the field of variables by examining their correlation coefficients. Following this, stepwise selection and maximum R^2 methods will be applied. The final best model or models will be verified by examining plots of the residuals and checking for multicollinearity.

IV. Data Analysis and Model Building

This chapter describes the results of the research methodology detailed in chapter III. First are discussions of data gathering and preparation. Data preparation includes determination of potential predictor variables, aggregation of the data, and exclusion of certain data. Following this is a correlational analysis, which results in elimination of variables that are highly correlated with others. Model generation and selection procedures are subsequently used to narrow the number of possible models to a small subset. Finally, the "best" model is selected, based upon an analysis of predictive power and verification of model assumptions.

Data Gathering

The Technical Information Center (TIC) at Tyndall AFB provided key information for finding base closure data, including points-of-contact in the Department of Defense's Base Closure Commission Office. Through this office, the authors were referred to Mr. Jeff Miller, who was very responsive to requests for data. Under the Freedom of Information Act, he was able to provide the 1993 closure data for Air Force active duty bases, ANG installations, AFRES installations, Air Force depots, and several Army installations. Miller was also able to provide a copy of the COBRA software program, and referred the authors to Major Charles Fletcher, who works on the contractual use of COBRA. Fletcher provided a computer diskette that contained the manual for the current COBRA software.

The COBRA data contained information on 34 Air Force (AF) installations, 14 Air Force Reserve (AFRES) installations, 10 Air National Guard (ANG) installations, 6 Air Force depots, and 8 Army installations. The COBRA data set included two closure scenarios for each of two Air Force installations, Shaw and Moody. In the case of the 8 Army installations, there were a total of 39 closure scenarios under consideration.

The COBRA software program has 13 input screens that are used to enter data for analysis. COBRA produces a report called INPUT.RPT, which contains the data entered from the input screens. Using COBRA to analyze the impact of each closure scenario, the INPUT.RPT reports were used to obtain the data for use in the regression procedure. Appendix C contains an illustrative INPUT.RPT file that pertains to Tyndall AFB.

Data Preparation

This section discusses the determination of potential predictor variables, aggregation of potential predictor variables, and exclusion of certain data records that seemed inconsistent.

Determination of Potential Regressors. As discussed in chapter III, the model requires as many as 250 inputs. In order to fulfill the goal of minimizing the number of variables in the CER, only those thought to be the most significant cost drivers were selected for further analysis. Those variables retained for further analysis included the following:

- Net present value (20 year)
- Distance to each base
- Mission equipment
- Support equipment
- Military light vehicles
- Heavy vehicles
- Special vehicles
- Officer transfer
- Total officers
- Enlisted transfer
- Total enlisted
- Civilian transfer
- Total civilians
- Percent of civilians not willing to move
- Base facilities
- Base acreage
- Freight cost
- Area cost factor
- Military light vehicle cost
- Heavy/special vehicle cost
- Non-payroll Real Property Maintenance Action costs (RPMA)
- Payroll RPMA costs
- Communications costs
- Non-payroll Base-Operations (base-ops) costs

- Payroll base-ops costs
- Military Family Housing (MFH) costs
- CHAMPUS inpatient costs
- CHAMPUS out-patient cost
- One-time moving costs
- Property transactions
- Construction avoidance
- MFH construction avoidance
- CHAMPUS inpatient visits
- CHAMPUS out-patient visits
- Military Construction costs (MILCON)

Some variables, such as number of officers transferred and construction avoidance, are time-phased over a six-year closure time period. In such cases, the six separate figures for each variable were summed across the six years to yield an aggregate figure.

Aggregation of Variables. After distilling to the listed set of potential cost drivers, many of them were combined to further reduce the set of variables to a manageable size. For example, the standard factors file used for the 1993 analysis set freight costs of mission and support equipment equal to each other. Consequently, they were combined to form one variable called *equipment*.

In other cases, variables were eliminated. For example, the number of personnel at the closing installation and the number of people transferred were rarely equal. COBRA documentation explained this as due to the elimination of manning positions. Despite the difference between the number of personnel assigned to an installation and the number that were assumed to be transferred, the two numbers tracked each other closely. Consequently, only the number of personnel at the closing installation was retained as a regressor variable.

The percent of civilians assumed unwilling to move was 10% for all AF, AFRES, and ANG installations; 60% for all depots, and 6% for all Army scenarios. Because this variable appeared to be a standard factor for each installation type, it was determined not to be a significant predictor of the response variable (20 year net present value, NPV).

COBRA initially divided RPMA and base operations costs into non-payroll and payroll variables, for a total of four variables. To reduce the number of required variables, non-payroll and payroll were summed together within each category, resulting in one RPMA cost and one base operations cost.

All construction costs and savings were combined to form an aggregate variable called *net construction:*

```
Net Construction = Total MILCON for the scenario
                   - total programmed construction avoided
                   - total programmed military family
                     housing construction avoided
```

Each of the variables on the right hand side of the equation is extracted from INPUT.RPT.

One final problem came in the treatment of the distances between the closing installation and gaining installations. From the several distances involved in nearly every closure scenario, it was deemed desirable to develop a single, unambiguous measure of distance. Table 4.1, an excerpt from the spreadsheet used in the data preparation process, shows examples of the multiple distances to the gaining installations. The possible solutions examined included using the average distance, a weighted distance based on equipment, a weighted distance based on personnel, a weighted distance based on vehicles, or a combination of the four weighted distances.

The person-weighted distance (***DI***) was used as a proxy for the distance traveled. Though it did not exactly mirror the distances moved by equipment or vehicles, it did provide a measure of the distance to all gaining installations. This variable was calculated using the following formula:

$$DI = \frac{\sum (dist.\ to\ gaining\ inst.)(\#\ personnel\ transferred)}{\sum (\#\ personnel\ transferred)}$$

TABLE 4.1

EXAMPLES OF DISTANCES TO GAINING INSTALLATIONS FROM EACH CLOSURE BASE.

Closure Base	Transfer Base	Distance to Base
Whiteman		
	McConnell	263
	Davis-Monthan	1271
	Ellsworth	769
	Base X	1000
	Columbia MO	87
Travis		
	Beale	92
	Base X	1000
Shaw II		
	Base X	1000
	Hill	2137
	Davis-Monthan	1957
	Mt Home	2402
	Pope	155
	Eglin	550

The following is a list of the resulting variables and aggregate variables. The list includes an acronym that will be used interchangeably with the respective variable in the remainder of the research. Appendix D provides the values of these variables in a base-by-base listing.

- NPV = Net present value (20 year)
- DI = Distance to base
- EQ = Equipment
- VH = Vehicles
- OF = Total officers
- EN = Total enlisted
- CI = Total civilians
- FA = Base facilities
- AC = Base acreage
- AR = Area cost factor
- RP = RPMA costs
- CO = Communications costs

* BO - Base-ops costs
* MF = MFH costs
* IP = CHAMPUS inpatient total costs
* OP = CHAMPUS out-patient total costs
* OT = One-time moving costs
* PT = Property transactions
* NC = Net construction

Inspection revealed that data associated with the various types of installations (e.g., depots, AFRES, etc.) have their own characteristics. Thus, dummy variables for ANG, AFRES, Depot, and Army installations were created.

* RV = Air Force Reserve installation
* NG = ANG installation
* DP = Air Force Depot
* MY = Army Installation

The dummy variables are meant to pick up costs associated with their respective types of installation. When all dummy variables are set equal to zero, the installation type defaults to active Air Force.

Data Exclusion. The Army data had some inconsistencies when compared to the Air Force data. The major difference was that the Army prevalently used multiple scenarios for each closing installation. The multiple scenarios differed from each other in many ways including number of gaining installations, number of people transferred, total personnel at the closing installation, and net construction.

Another difference from the Air Force data was that the Army inputs for the vehicles variables were given in tons of vehicles rather than number of vehicles. Also, the

Presidio, FSTC, and ITAC had inputs of either a one or zero for square footage of facilities. Additionally, none of the Army installations had any inputs for CHAMPUS costs.

Because of the many differences from the rest of the data set, the Army data was eliminated from the final results.

Correlational Analysis

A correlation analysis was used to identify variables that would cause collinearity during the multivariate regression. The correlation analysis was completed on the remaining data set -- active duty Air Force, AFRES, ANG, and Air Force depot installations. For simplicity, this data set will hereafter be referred to as the Air Force data.

The correlation analysis of the Air Force data, given in Appendix E, revealed strong correlations. Table 4.2 displays all correlation coefficients greater than 0.80. A correlation coefficient of 0.80 is the bottom limit of the strong relationship range, 0.80 to 1.00 (Devore, 1991:205).

To reduce the possibility of collinearity in the model generation and selection procedures, elimination of all strong correlation relationships is important. Simple techniques include elimination or aggregation of variables. The sources of the correlations of Table 4.2 may be subdivided into two categories -- personnel-related and

TABLE 4.2

VARIABLES WITH STRONG CORRELATION COEFFICIENTS (AIR FORCE).

Variables		Correlation Coefficient
CI	FA	0.937
OT	DP	0.929
RP	FA	0.895
CI	DP	0.889
OF	EN	0.888
CI	OT	0.885
FA	OT	0.825
CI	RP	0.807

facilities-related. These are the subjects of the next sections.

Personnel-Related (CI-FA, CI-DP, OF-EN, CI-OT, CI-RP). Air Force guidance (AF Regulation 85-series) gives the amount of facility square footage that is considered adequate for different operations and number of personnel. The result is that the facilities variable is based on the number of personnel and the mission. This is substantiated by reviewing the correlation analysis in Appendix E, which shows a strong correlation between total civilians and square footage of facilities (0.937) and moderate correlations between facilities and the number officers (0.574) and enlisted personnel (0.546). The facilities variable is also highly correlated with non-personnel variables and will be further examined in the next section, facilities-related correlations.

The numbers of officers and enlisted (OF-EN) were highly correlated (0.888). Because the effects of both on the response variable were assumed to be important, they were combined to create a new variable called military personnel (MILPER). Another variable was created called total personnel (TOTPER). It was created by summing the total number of officers, enlisted, and civilians at an installation.

A correlational test of the Air Force data was accomplished using combinations of the newly created personnel variables and the former variables: (1) OF, EN, and CI as separate variables; (2) MILPER and CI; and (3) TOTPER. The results were that combination (1) produced many strong correlations, combination (2) had slightly fewer highly correlated variables, and combination (3) had strong correlations between only two variables, TOTPER and RP. Because of the greatly reduced number of strong correlations from using combination (3), TOTPER was chosen to replace OF, EN, and CI in the Air Force data set.

Facilities-Related (OT-DP, RP-FA, FA-OT). The values of OT and DP were reviewed because of the strong correlation between the one time moving costs variable and the depot dummy variable. A review of the data set showed that high one time moving costs were consistent only among depots.

One time moving cost was eliminated from the list of variables under consideration because the depot dummy variable could capture the effect of the one time moving cost. Inclusion of the dummy variable rather than the one time cost also requires less time investment by the user.

The authors' experience at base-level Civil Engineering suggests that Real Property Maintenance costs (RPMA) are a direct result of performing maintenance and repair on the facilities (FA). Therefore, it is not surprising that the variables are highly correlated. As mentioned in the previous section, square footage of facilities is also correlated to the number of personnel. The effects of the facilities variable could be captured by other variables and was, therefore, excluded from further analysis.

Summary of Correlation Analyses. The variables remaining after elimination or aggregation of those with strong correlations were RP, DP, and TOTPER. The subsequent correlation analysis revealed that TOTPER was still strongly correlated with RP (0.8823) and moderately correlated with DP (0.7355). Because each was presumed to have significance as a possible predictor, however, the correlations were tolerated and the variables were retained for the regression analysis.

Model Generation and Selection

The SAS software program was used in the model generation procedure. The SAS statistical procedures are controlled by an input instruction file that the user creates. For this analysis, the input program told SAS to open the data file, read in all the variables, create models using the Stepwise and Maximum R^2 procedures, and output the results to a text file. The SAS instruction files are included as Appendix F.

Control of the Stepwise and Maximum R^2 procedures enabled the authors to greatly narrow the number of models. Specifically, variable entry and exit controls in the Stepwise procedure made it possible to reduce the number of models reviewed only to those whose parameters were significant to the 90 percent confidence level or better (according to the t-test). Additionally, the SAS program allows output of the residuals, predicted values, and variance inflation factors.

Model Selection. A review of the output was done to rank the models by their respective R^2 values and to ensure that the individual parameters of each model were significant to better than the $\alpha=0.10$ level (90% confidence). The top six models, their parametric statistics, and the accompanying ANOVA tables are displayed below.

Model 1.

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION OF NPV
NOTE: MODEL FORCED THROUGH ORIGIN

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P	VIF
DI	-86275.6	30351.1	-2.84	0.0063	0.9
EQ	35786.6	14920.6	2.40	0.0199	0.7
VH	-1.569E+05	41156.2	-3.81	0.0004	1.6
AC	-4020.41	1683.56	-2.39	0.0205	1.3
RP	-8.66145	2.93463	-2.95	0.0047	3.8
BO	-24.7609	5.41280	-4.57	0.0000	2.2
MF	-20.8032	7.13171	-2.92	0.0051	2.7
IP	8.95854	3.29968	2.71	0.0089	1.7
NC	1.10306	0.08974	12.29	0.0000	2.4
DP	4.065E+08	7.190E+07	5.65	0.0000	3.4

R-SQUARED	0.9133	RESID. MEAN SQUARE (MSE)	8.346E+15
ADJ R-SQUARED	0.8972	STANDARD DEVIATION	9.136E+07

SOURCE	DF	SS	MS	F	P
REGRESSION	10	4.746E+18	4.746E+17	56.87	0.0000
RESIDUAL	54	4.507E+17	8.346E+15		
TOTAL	63	5.197E+18			

CASES INCLUDED 64 MISSING CASES 0

Model 2

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION OF NPV
NOTE: MODEL FORCED THROUGH ORIGIN

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P	VIF
VH	-1.478E+05	39356.8	-3.75	0.0004	1.4
AC	-4167.63	1768.14	-2.36	0.0219	1.3
RP	-6.30637	2.58563	-2.44	0.0179	2.6
BO	-22.1124	5.28164	-4.19	0.0001	1.9
MF	-27.5226	6.88293	-4.00	0.0002	2.3
IP	6.43575	3.34768	1.92	0.0596	1.6
NC	1.09118	0.08909	12.25	0.0000	2.1
DP	3.226E+08	6.643E+07	4.86	0.0000	2.6

R-SQUARED	0.9002	RESID. MEAN SQUARE (MSE)	9.259E+15
ADJ R-SQUARED	0.8860	STANDARD DEVIATION	9.623E+07

SOURCE	DF	SS	MS	F	P
REGRESSION	8	4.679E+18	5.848E+17	63.16	0.0000
RESIDUAL	56	5.185E+17	9.259E+15		
TOTAL	63	5.197E+18			

CASES INCLUDED 64 MISSING CASES 0

Model 3.

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION OF NPV
NOTE: MODEL FORCED THROUGH ORIGIN

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P	VIF
VH	-1.487E+05	40273.7	-3.69	0.0005	1.4
AC	-4085.25	1808.94	-2.26	0.0278	1.3
RP	-7.18406	2.60450	-2.76	0.0078	2.6
BO	-22.4882	5.40140	-4.16	0.0001	1.9
MF	-27.5997	7.04369	-3.92	0.0002	2.3
NC	1.04070	0.08712	11.94	0.0000	1.9
DP	3.869E+08	5.872E+07	6.59	0.0000	1.9

R-SQUARED	0.8936	RESID. MEAN SQUARE (MSE)	9.697E+15
ADJ R-SQUARED	0.8806	STANDARD DEVIATION	9.847E+07

SOURCE	DF	SS	MS	F	P
REGRESSION	7	4.644E+18	6.635E+17	68.42	0.0000
RESIDUAL	57	5.527E+17	9.697E+15		
TOTAL	63	5.197E+18			

CASES INCLUDED 64 MISSING CASES 0

Model 4.

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION OF NPV

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P	VIF
CONSTANT	-2.560E+08	4.811E+07	-5.32	0.0000	
VH	-88160.0	39785.3	-2.22	0.0309	1.7
AC	-3736.66	1583.65	-2.36	0.0219	1.3
RP	-8.21521	2.43209	-3.38	0.0013	2.9
BO	-12.6998	6.00250	-2.12	0.0389	3.1
NC	1.08538	0.08040	13.50	0.0000	2.1
RV	2.644E+08	4.677E+07	5.65	0.0000	3.2
NG	2.253E+08	4.115E+07	5.48	0.0000	1.9
DP	4.856E+08	5.109E+07	9.50	0.0000	1.9

R-SQUARED	0.8873	RESID. MEAN SQUARE (MSE)	7.419E+15
ADJ R-SQUARED	0.8709	STANDARD DEVIATION	8.613E+07

SOURCE	DF	SS	MS	F	P
REGRESSION	8	3.211E+18	4.014E+17	54.11	0.0000
RESIDUAL	55	4.080E+17	7.419E+15		
TOTAL	63	3.619E+18			

CASES INCLUDED 64 MISSING CASES 0

Model 5.

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION OF NPV

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P	VIF
CONSTANT	-1.514E+08	6.498E+07	-2.33	0.0236	
EQ	51490.2	23164.8	2.22	0.0304	1.6
VH	-1.575E+05	42098.6	-3.74	0.0004	1.6
AC	-4079.46	1720.28	-2.37	0.0213	1.3
RP	-7.74683	2.94808	-2.63	0.0112	3.7
BO	-21.8226	5.48140	-3.98	0.0002	2.2
MF	-24.3701	7.02535	-3.47	0.0010	2.5
IP	8.22408	3.33763	2.46	0.0170	1.7
NC	1.13475	0.09307	12.19	0.0000	2.4
DP	3.768E+08	7.100E+07	5.31	0.0000	3.1

R-SQUARED	0.8699	RESID. MEAN SQUARE (MSE)	8.718E+15
ADJ R-SQUARED	0.8482	STANDARD DEVIATION	9.337E+07

SOURCE	DF	SS	MS	F	P
REGRESSION	9	3.148E+18	3.498E+17	40.12	0.0000
RESIDUAL	54	4.708E+17	8.718E+15		
TOTAL	63	3.619E+18			

CASES INCLUDED 64   MISSING CASES 0

Model 6.

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION OF NPV

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P	VIF
CONSTANT	-7.855E+07	4.038E+07	-1.95	0.0569	
VH	-1.321E+05	42466.4	-3.11	0.0030	1.6
AC	-4166.50	1775.55	-2.35	0.0226	1.3
RP	-6.97026	2.81076	-2.48	0.0162	3.1
BO	-13.7491	6.72099	-2.05	0.0456	3.1
MF	-25.6533	7.12035	-3.60	0.0007	2.5
NC	1.02002	0.08854	11.52	0.0000	2.1
RV	9.030E+07	4.382E+07	2.06	0.0441	2.3
DP	3.979E+08	5.864E+07	6.79	0.0000	2.0

R-SQUARED	0.8591	RESID. MEAN SQUARE (MSE)	9.274E+15
ADJ R-SQUARED	0.8386	STANDARD DEVIATION	9.630E+07

SOURCE	DF	SS	MS	F	P
REGRESSION	8	3.109E+18	3.886E+17	41.91	0.0000
RESIDUAL	55	5.101E+17	9.274E+15		
TOTAL	63	3.619E+18			

CASES INCLUDED 64   MISSING CASES 0

The model with the highest R^2 value is Model 1, which contains the variables DI, EQ, VH, AC, RP, BO, MF, IP, NC, and DP. It has an R^2 value of 0.9133 (i.e., the model explains 91.33% of the variance).

Gauss-Markov Assumptions. Because Model 1 had the highest R^2 value, it was checked first to determine if the Gauss-Markov assumptions were met. The tests for adequacy, from chapter III, include tests for multicollinearity, non-normality of the residuals, and heteroscedasticity.

Multicollinearity. A review of the VIFs for this model, listed in the output above, shows that all VIF values were below the value of 5.0 established in chapter III. This shows the model does not suffer from multicollinearity.

Normality of the Residuals. This test is accomplished by creating a histogram of the studentized residuals. If the resulting distribution appears to approximate a normal distribution, then the assumption holds. An examination of the histogram for Model 1, displayed in Figure 4.1 below, shows that the distribution is approximately normal.

If a plot of the residuals shows departure from a normal distribution, there could be a problem using the least squares method of regression. However, the most critical departure from normal is "heaviness" in the tails of the distribution. Heaviness in the tails indicates significant outliers which could tend to "pull" the

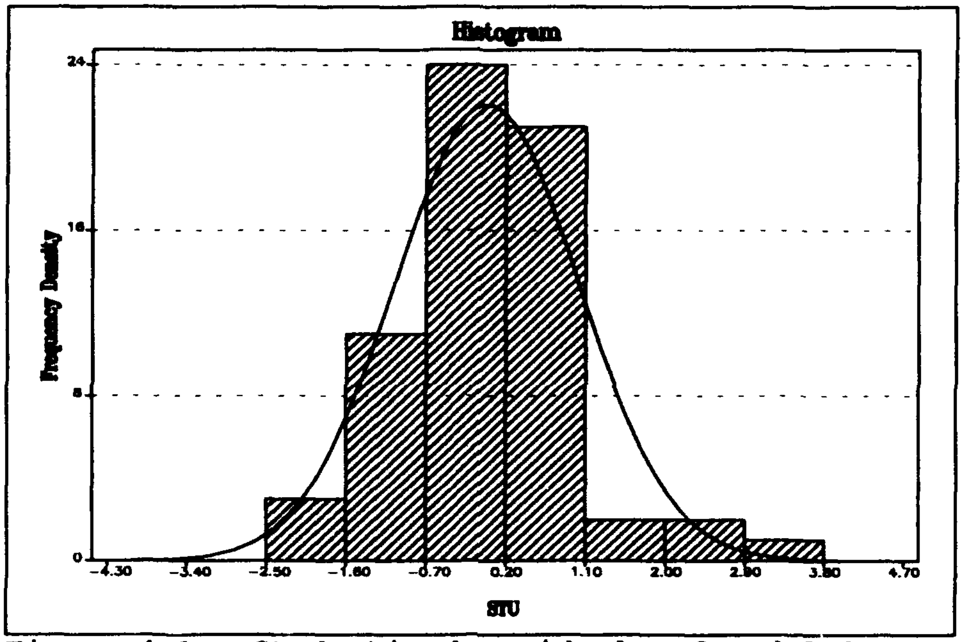


Figure 4.1. Studentized residuals of Model 1.

regression in their direction (Montgomery & Peck, 1982;364). Examination of the plot of residuals for Model 1 does show some departure from normality; however, the departure is mainly focused near zero standard deviations. Heavy outliers (heaviness in the tails) do not appear to be a problem.

Nonetheless, the slight deviations from normal did cause some apprehension. An examination was accomplished of the histograms of the studentized residuals for the top six models. All models seemed to follow this trend of deviation from normal in the -0.5 to 0.5 standard deviations range (see Appendices G-L for histograms of the top six models). However, according to Wesolowsky,

> Given that the error term ε has constant variance and ε_i's are statistically independent, deviations from normality in the distribution of ε usually have no serious consequences. (Wesolowsky, 1976:125)

This concept is also supported by Theil, who states that for the F and t tests to be affected, the histogram's values must differ *considerably* from the normal density distribution (Theil, 1971;616). Thus, non-normality probably is not a problem in the case of Model 1.

Test for Heteroscedasticity. To determine if the residuals have a non-constant variance, the residuals are plotted against the predicted values. A plot of the residuals against the predicted values for Model 1 is shown in Figure 4.2. The plot shows no discernible indication of a trend. Consequently, the assumption of constant variance is considered valid.

The scatter plots of the residuals against each of the predictor variables for Model 1 are included in Appendix G. With the exception of DP, none of the variables show an apparent trend in the residual plots. An inspection of the plot of residuals against DP shows that the variance of the residuals at the value of 1 (the dummy variable designation) appears to be smaller than for the value of 0. This can be explained, however, by the fact that there are only six plots at the value of 1 (depot designators) and fifty-eight

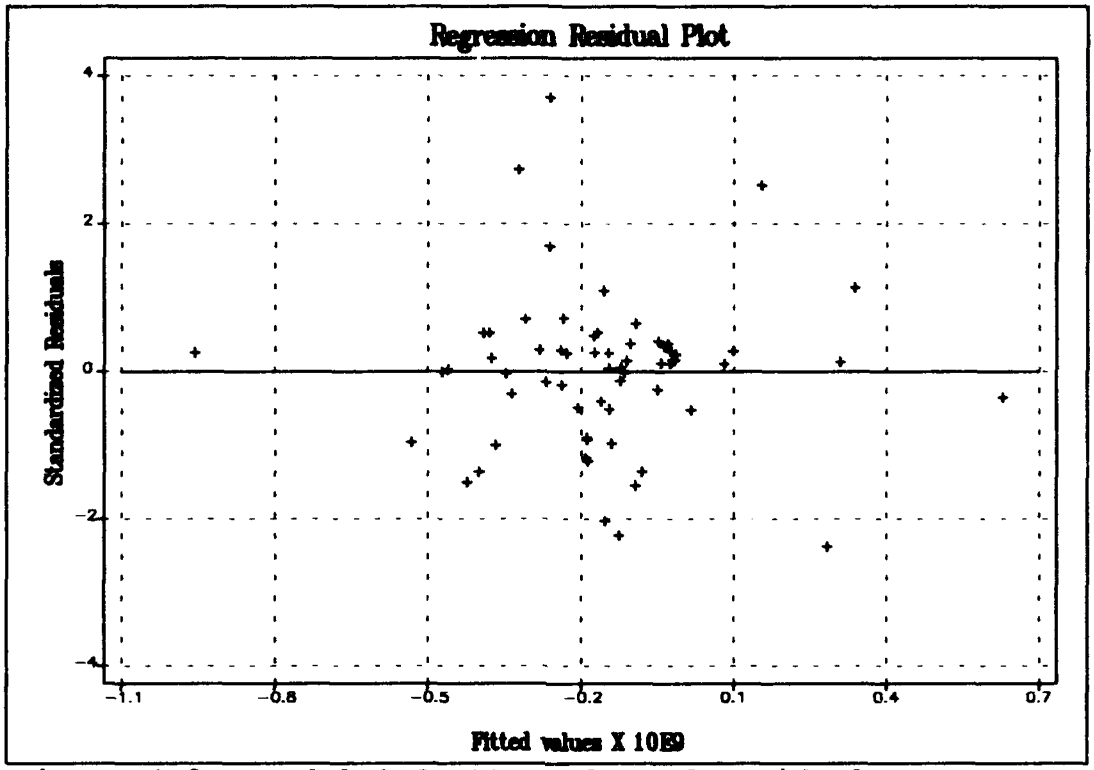


Figure 4.2. Model 1 Scatter Plot of residuals versus predicted values.

at the value of 0. Again, the assumption of constant variance in the residuals appears to be valid.

<u>Selection of the Best Model</u>. The Gauss-Markov assumptions for zero correlation between residuals and for homoscedasticity both appear to be valid for Model 1. The distribution of residuals was closely normal. Additionally, the plot of the residuals against the independent variables showed no significant trends.

Model 1 appears to satisfy the Gauss-Markov assumptions. Therefore, it is selected as the best model for predicting base closure costs. The selected model is given by

$$\begin{aligned}\mathbf{NPV} = &-86{,}275.6(\mathbf{DI}) + 35{,}786.6(\mathbf{EQ}) - 156{,}900(\mathbf{VH}) - 4{,}020.41(\mathbf{AC}) \\ &- 8.66145(\mathbf{RP}) - 24.7609(\mathbf{BO}) - 20.8032(\mathbf{MF}) + 8.95854(\mathbf{IP}) \\ &+ 1.10306(\mathbf{NC}) + 406{,}500{,}000(\mathbf{DP})\end{aligned}$$

Definitions of the variables used in the selected model, Model 1, are provided in Table 4.3. Explanations of the aggregate variables are given as well as COBRA definitions of the root variables.

TABLE 4.3

DEFINITIONS OF VARIABLES IN SELECTED MODEL

NPV	20-year Net Present Value (dollars)	This value is given in COBSUM.RPT, Realignment Summary Report. This is a measure of the total costs (over the 20-year period of analysis) to be realized by taking the closure/realignment actions in the scenario. The more negative the number, the greater the savings. A positive number actually represents a cost over the 20-year period
DI	Person-Distance (miles)	The "weighted" distance of all personnel transferred to gaining installations. The distances to each gaining installation are requested in input screen 2. The number of personnel transferred to each gaining installation is requested in input screen 3. The number of personnel transferred is given by officers, enlisted, and civilians. A sum of total personnel moving to each installation is taken to calculate the weighted distance according to the formula given in chapter IV.
RP	Real Property Maintenance Activities (dollars)	This is an aggregate of the RPMA payroll and non-payroll budgets. These values are requested in input screen 4. Each is the respective budget for the base at the beginning of the scenario (does not include Military Family, MFH, costs).

TABLE 4.3 (continued)

AC	Acreage	This is the total acres on the base at the beginning of the scenario.
EQ	Equipment (Tons)	This variable is an aggregate of the tons of mission equipment and the tons of support equipment. Both these values are requested in input screen 3.
VH	Vehicles (number of vehicles)	Vehicles is an aggregate of the number of military light vehicles and the number of heavy/special vehicles. These variables are requested in input screen 3. Military light vehicles is the number of vehicles which will be driven from the closing installation to each of the gaining installations. Heavy/special vehicles are those that will be transported to gaining installations.
BO	Base Operations (dollars)	This is an aggregate of the Base Operations payroll and non-payroll budgets. These values are requested in input screen 4. Each is the respective budget for the base at the beginning of the scenario.
MF	Family Housing Costs ($)	This variable is the total family housing budget at the beginning of the scenario. It is requested in input screen 4.
IP	CHAMPUS In-Patient Costs per year (dollars)	This variable is formed by multiplying the $/visit for CHAMPUS in-patients by the number of visits per year. The $/visit is requested in input screen 4. The number of visits per year is requested in input screen 6. The number of visits per year is actually the yearly (summed over the six year closure scenario) change in the number of in-patient visits of retirees and their dependents to the on-base hospital/treatment facilities.
NC	Net Construction (dollars)	This variable is an aggregate of the total Military Construction (MILCON) to be completed because of the scenario closure action (at gaining and closing installations) **minus** the construction costs avoided at the closing installation **minus** family housing construction costs avoided at the closing installation. MILCON activities are requested in input screen 7. Construction avoidance is requested in input screen 5. Family housing construction avoided is requested in input screen 5.
DP	Depot	This is a dummy variable used to designate the closure installation as a depot. Its entry is either a 1 (for depots) or 0 (for other)

Model Validation

This model has passed the test for adequacy; however, prior to being implemented in its suggested use, it should be validated. Three possible methods of model validation are discussed here.

Data Splitting. Because of the limited sample size, all values were used in the regression. Setting aside some of the data for testing predictive performance would have created a smaller sample size. A smaller data set would, however, have reduced the degrees of freedom in the regression analysis, thereby making the desired confidence levels more difficult to achieve.

Collection of Fresh Data. If another round of base closures is contemplated, fresh data (estimations) would be available from the BRAC Commission and could be used to validate the parametric model. Moreover, installations have been closing for several years. Collection of cost data from these closures would provide a direct comparison of this model to reality.

Comparison of Forecasted Results. Using the original data, the predicted net present values (NPV) from COBRA and the selected model were compared. Table 4.4 gives both models' predicted net present values. The installations in the table are ranked in ascending order of cost (negative

TABLE 4.4
PREDICTED COSTS AND RANKINGS OF COBRA VERSUS THE PARAMETRIC MODEL
[parentheses indicate negative costs]

BASE	COBRA NPV	Rank	Rank	Parametric NPV
Homestead	($940,268,000)	1	1	($957,825,579)
Grand_Forks	($610,729,000)	2	2	($531,013,039)
Plattsburgh	($470,858,000)	5	3	($470,098,347)
Shaw_I	($456,926,000)	6	4	($459,150,840)
McGuire	($544,376,000)	3	5	($421,430,289)
Minot	($511,371,000)	4	6	($399,769,636)
Beale	($347,585,000)	11	7	($390,509,437)
Tyndall	($337,166,000)	12	8	($379,062,704)
Davis_Monthan	($360,473,000)	8	9	($375,437,672)
Moody_I	($453,257,000)	7	10	($365,691,444)
K_I_Sawyer	($348,897,000)	10	11	($345,992,255)
Holloman	($353,603,000)	9	12	($336,175,322)
Whiteman	($85,447,000)	42	13	($321,623,006)
March	($247,492,000)	23	14	($308,770,012)
Shaw_II	($256,303,000)	20	15	($281,389,907)
Seymour_Johnson	($280,000,000)	17	16	($266,919,995)
Fairchild	($113,853,000)	40	17	($261,796,932)
Malmstrom	$56,950,000	57	18	($258,749,429)
Charleston	($212,911,000)	26	19	($238,273,006)
Grissom	($253,222,000)	21	20	($237,480,260)
OHare	($172,149,000)	31	21	($233,677,621)
Carswell	($205,872,000)	27	22	($226,853,039)
Moody_II	($248,356,000)	22	23	($205,455,738)
Griffiss	($295,612,000)	15	24	($191,141,485)
Altus	($269,672,000)	18	25	($189,041,317)
Selfridge	($260,373,000)	19	26	($188,072,504)
McConnell	($295,355,000)	16	27	($187,470,069)
General_Mitchell	($132,080,000)	35	28	($174,297,108)
Barksdale	($151,376,000)	32	29	($172,951,867)
Greater_Pittsburgh	($120,133,000)	37	30	($165,678,680)
Bergstrom	($196,499,000)	29	31	($160,519,977)
Fresno	($61,211,000)	45	32	($155,410,054)
Ellsworth	($324,505,000)	13	33	($151,204,234)
Cannon	($188,682,000)	30	34	($145,026,328)
Youngstown	($122,098,000)	36	35	($144,546,939)

TABLE 4.4 (continued)

BASE	COBRA NPV	Rank	Rank	Parametric NPV
Dover	($140,220,000)	33	36	($143,320,509)
Mountain_Home	($224,988,000)	24	37	($138,553,902)
Pope	($314,859,000)	14	38	($126,499,323)
McChord	($133,457,000)	34	39	($122,897,042)
Buckley	($117,467,000)	38	40	($120,763,502)
Dobbins	($116,153,000)	39	41	($114,016,308)
Minn_StPaul	($97,372,000)	41	42	($110,377,093)
Portland	($70,374,000)	43	43	($103,662,513)
Luke	($34,325,000)	46	44	($91,538,971)
Little_Rock	($222,361,000)	25	45	($91,125,903)
Otis	($201,678,000)	28	46	($81,494,166)
McEntire	($13,054,000)	50	47	($48,438,639)
Westover	($69,685,000)	44	48	($47,953,977)
Ellington	($34,205,000)	47	49	($44,179,285)
Stewart	($7,418,000)	51	50	($37,703,426)
Tucson	$4,171,000	55	51	($29,247,161)
Boise	($16,075,000)	49	52	($25,143,419)
Salt_Lake_City	($313,000)	54	53	($23,871,166)
Great_Falls	($6,556,000)	52	54	($20,141,590)
Martin_State	($2,777,000)	53	55	($15,793,999)
Pittsburgh	$5,481,000	56	56	($15,047,029)
McClellan	($25,363,000)	48	57	$15,846,152
Dyess	$87,703,000	58	58	$80,441,039
Travis	$117,950,000	59	59	$99,484,883
Tinker	$278,587,000	61	60	$155,052,355
Newark	$126,802,000	60	61	$283,580,075
Robins	$315,695,000	62	62	$308,315,085
Hill	$424,725,000	63	63	$337,428,739
Kelly	$610,033,000	64	64	$630,256,594

costs are benefits). Both rankings have the same installations listed within the top twelve. After the twelfth position, the differences in the NPVs are relatively small, and any slight deviations in the predicted values can significantly change the order of the rankings. A few outliers such as Whiteman and Malmstrom had unique closure activities that are not captured by the parametric model and, consequently, have notable differences in predicted costs.

The Spearman's Correlation Test was used to determine the statistical similarity between the ranked listings. This test checks to see if a rise or fall in the value of one list is associated with a corresponding rise or fall of the other -- the existence of direct correlation (Langley, 1971:199).

To test for correlation, the null hypothesis is that there is no significant correlation between the two ranked listings. Next, the "probability of getting, by chance, a difference between the rank values as small, or as large [for inverse correlation], as that which is observed" is determined (Langley, 1971:201). If the probability is remote, according to a selected confidence level, the null hypothesis is rejected.

An outline of the test is as follows:

1.) Calculate the absolute difference between each rank value and its paired partner, d.

2.) Square the difference, d^2.

3.) Sum the squares of the differences, D^2.

4.) Use the following formula to calculate z (n is the number of observations in the ranked lists):

$$z = \sqrt{n-1} * \left| 1 - \frac{D^2}{(n^3 - n)/6} \right|$$

5.) The value of z can be used on a cumulative standard normal table to find the critical probability value.

The result of the similarity test for the rankings given in Table 4.4 is to reject the null hypothesis:

H_0: There is no direct correlation between rank orders of the two models.

D^2 = 5,902 n = 64 z = 6.865

According to the determined value of z, the critical probability is at a significance level of $\alpha << 0.01$. Therefore, reject H_0. The ranked listings are statistically similar.

V. Results and Summary

The Base Realignment and Closure Commission is responsible for selecting installations for possible closure. Selection is a two-phase process taking into account both military value and economic benefit. BRAC uses the Cost of Base Realignment Action (COBRA) model to predict the economic benefits of closure actions.

COBRA is an economic analysis model which requires extensive estimation and data collection on the part of the user. It provides detailed reports of the costs and benefits of closure and realignment scenario. With the many installations contemplated for each round of closure, verifying the data and performing quality checks of the inputs can be an enormous task. This thesis developed a parametric cost model designed to capture the essential aspects of COBRA, requiring fewer input values and less data collection.

Data from the 1993 base closure analysis was provided for model development. This data included the COBRA inputs for 64 Air Force installations (active duty, reserve, national guard, and depot). Data aggregation, inspection, and preparation narrowed the variable field to 18 potential cost drivers. Correlational analysis was completed which provided justification to narrow the field even further.

Once the final set of variables was determined, the stepwise selection and maximum R^2 techniques were applied to the data. These widely accepted methods of variable selection produced several models for final analysis. The authors chose the top six models according to their respective R^2 statistic (ability to explain variance). Starting with the top model, the residuals were examined to check for normality in the distribution. Scatter plots of the residuals were also examined to verify that heteroscedasticity did not exist. The variance inflation factors for each variable in the model were examined as a check for multicollinearity.

Model 1 passed all these tests, thereby verifying its adequacy. Because this model explained the most variance (highest R^2 value) and passed the model adequacy test, further analysis of the other models was not required. The selected model uses only ten variables (though some were aggregates) and explained over 91 percent of the variance.

The concluding discussion in chapter IV showed that the COBRA and parametric models produce statistically similar rankings over the entire data set. The BRAC Commission, however, may decide to divide the installations into groups (active duty, AFRES, ANG, depot) and rank them within the particular groups. Without knowing the Commission's closure goals or criteria for grouping, the authors were not able to test the model by ranking installations divided into different groups.

The model developed in this thesis did fulfill the goal of capturing the essential aspects of COBRA while minimizing the number of inputs. In the sensitive environment of base closure, the detailed analysis and output of COBRA may be necessary. However, the sensitivity also requires accuracy in the data. A parametric model, such as the one developed in this thesis, can be used for a "quick" analysis of the entire field of contemplated closure installations. With the few number of variables necessary for prediction, the data can be checked for quality and consistency. The field of potential closure bases can then be narrowed to only those requiring detailed study by COBRA.

Additional Remarks

Throughout the available literature, stepwise selection appeared to be the preferred model selection procedure, mainly because of its rigorous check of variable combinations. In many texts, however, suggestions were made to use an additional procedure in conjunction with stepwise. This thesis employed both the stepwise and maximum R^2 procedures. Using two procedures proved to be worthwhile -- the selected model was produced by the maximum R^2 procedure.

One cautionary measure is that use of the parameter estimates and the respective variables is only valid when used in conjunction with the entire model statement. No variable can be pulled out of the equation and used with its accompanying parameter to calculate a cost or savings. Each parameter was estimated with interrelations to the other parameters and variables. Also, estimates of the variables used in the model must be made in the same manner as if they were going to be used in the COBRA model.

Also, the estimates from the parametric model were made using the exact parameters as they were calculated by SAS. Examination of the standard error for the parameters shows that each is significant to only one or two digits. For further validation and use, these parameters should be rounded to the appropriate significant digits.

The same is also true for the comparison of estimates between COBRA and the parametric model. The standard deviation for the selected model is 9.136E+07. Therefore, comparisons of the NPV estimates should only be made at two or three significant digits. These round-offs would only improve the statistical similarity of the compared listings.

Recommendations

This thesis was one attempt at simplifying the cost estimating required in the BRAC process. The model appears to be satisfactory; however, no one will ever know which parametric model is "best." Perhaps other COBRA variables (or aggregates) not considered in this thesis would lead to a better model; even a nonlinear function may provide a better fit to the data. One could always include more variables and get a better prediction, but the initial goal of developing the model must be kept in mind -- to explain the most with the least.

In the BRAC 93 standard factors file used to calculate the COBRA estimates, the discount rate was set at seven percent. Should an alternate discount rate be chosen, a simple change can be made to the standard factors file and COBRA will give appropriate estimates. The coefficients of the parametric model, however, are a function of the seven percent rate; therefore, its estimates are only valid at this rate. A recommendation for further work is to complete a sensitivity analysis of the estimated NPVs upon the discount rate. If a factor could be included in the model to account for the discount rate, the model would be more flexible to changing rates.

In any model, validation is important prior to implementation in the field. Though the model developed in this thesis was shown to be statistically adequate, it has only been shown to be valid according to one test. The techniques of model validation (analysis of parameter coefficients, data splitting, and gathering fresh data) discussed in chapter III should be applied, if possible. The authors recommend collecting more data (assuming another round of base closure will take place) for validation. The data can be split with some being used to repeat the procedures described in this thesis. Replicating the model development procedures will enable the researcher to "fine tune" the parameter estimates. The remaining data can be used to test the predictive capability of the model versus COBRA.

An alternative validation test would be to collect actual cost data from recent and ongoing installation closures. These actual costs can then be compared to both the COBRA and parametric estimates. This would provide a comparison of both models to reality.

APPENDICES

Appendix A: Installation Task Forces and Categories

(Defense Secretary, 1988;49)

TASK FORCE	CATEGORY
Ground	Operating Ground Operating Troops
Air	Operating Tactical Aircraft Operating Strategic Aircraft Operating Mobility Aircraft Operating Missiles Flying Training
Sea	Operating Surface Ships Operating Submarines
Training and Administration	Headquarters Training Classrooms
Depot	Maintenance Depots Supply Depots Munitions Facilities Industrial Facilities Productions Facilities
All Other	Guard & Reserve Centers Communications/Intelligence Sites R&D Laboratories Special Operations Bases Space Operations Centers Medical Facilities

Appendix B: Military Value Factors and Physical Attributes

(Defense Secretary, 1988;50)

FACTORS	PHYSICAL ATTRIBUTES
Mission Suitability	Site-Specific Mission Deployment Means Relationship to Other Activities Weather/Terrain/Land Use Survivability Maneuver Space
Availability of Facilities	Operations Support Infrastructure Administration
Quality of Facilities	Condition Technology Configuration
Quality of Life	Family Housing Bachelor Housing Recreation/Amenities Medical
Community Support	Work Force Commercial Transport Infrastructure Complementary Industry

Appendix C: Example of COBRA Input Report

INPUT SCREEN ONE - GENERAL SCENARIO (COBRA v4.04)
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Group : Small Aircraft
Service : USAF
Option Package : Tyndall

Model Year One : FY 1994

Model does Time-Phasing of Construction/Shutdown: No

Base Name	Strategy:
Tyndall, FL	Closes in 1995
Shaw, SC	Realignment
Barksdale, LA	Realignment
Eglin, FL	Realignment
Base X Realignment	

Summary:
Close Tyndall FY95/4; F15s move to Shaw; Drones move to Eglin; AFCE moves to Barksdale; 1 AF and SE Air Defense move to Eglin; Weapons controller school moves to Eglin

Revised CHAMPUS
File Name: 8Tyndall

INPUT SCREEN TWO - DISTANCE TABLE (COBRA v4.04)
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

From Base:	To Base:	Distance:
Tyndall, FL	Shaw, SC	531.0 mi
Tyndall, FL	Barksdale, LA	577.0 mi
Tyndall, FL	Eglin, FL	82.0 mi
Tyndall, FL	Base X	1,000.0 mi
Shaw, SC	Barksdale, LA	1,000.0 mi
Shaw, SC	Eglin, FL	1,000.0 mi
Shaw, SC	Base X	1,000.0 mi
Barksdale, LA	Eglin, FL	1,000.0 mi
Barksdale, LA	Base X	1,000.0 mi
Eglin, FL	Base X	1,000.0 mi

INPUT SCREEN THREE - MOVEMENT TABLE (COBRA v4.04) - Page 2
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Transfers from Tyndall, FL to Shaw, SC

	1994	1995	1996	1997	1998	1999
Officers:	0	147	0	0	0	0
Enlisted:	0	1,602	0	0	0	0
Civilians:	0	54	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	1,000	0	0	0	0
Suppt Eqpt (tons):	0	500	0	0	0	0
Mil Light Vehic:	0	160	0	0	0	0
Heavy/Spec Vehic:	0	115	0	0	0	0

Transfers from Shaw, SC to Tyndall, FL

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

INPUT SCREEN THREE - MOVEMENT TABLE (COBRA v4.04) - Page 3
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Transfers from Tyndall, FL to Barksdale, LA

	1994	1995	1996	1997	1998	1999
Officers:	0	57	0	0	0	0
Enlisted:	0	91	0	0	0	0
Civilians:	0	155	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

Transfers from Barksdale, LA to Tyndall, FL

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

INPUT SCREEN THREE - MOVEMENT TABLE (COBRA v4.04) - Page 4
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Transfers from Tyndall, FL to Eglin, FL

	1994	1995	1996	1997	1998	1999
Officers:	0	356	0	0	0	0
Enlisted:	0	858	0	0	0	0
Civilians:	0	244	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	1,000	0	0	0	0
Suppt Eqpt (tons):	0	500	0	0	0	0
Mil Light Vehic:	0	160	0	0	0	0
Heavy/Spec Vehic:	0	115	0	0	0	0

Transfers from Eglin, FL to Tyndall, FL

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

INPUT SCREEN THREE - MOVEMENT TABLE (COBRA v4.04) - Page 5
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Transfers from Tyndall, FL to Base X

	1994	1995	1996	1997	1998	1999
Officers:	0	141	0	0	0	0
Enlisted:	0	378	0	0	0	0
Civilians:	0	137	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

Transfers from Base X to Tyndall, FL

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons).	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

INPUT SCREEN THREE - MOVEMENT TABLE (COBRA v4.04) - Page 6
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Transfers from Shaw, SC to Barksdale, LA

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

Transfers from Barksdale, LA to Shaw, SC

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

INPUT SCREEN THREE - MOVEMENT TABLE (COBRA v4.04) - Page 7
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Transfers from Shaw, SC to Eglin, FL

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

Transfers from Eglin, FL to Shaw, SC

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

INPUT SCREEN THREE - MOVEMENT TABLE (COBRA v4.04) - Page 8
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Transfers from Shaw, SC to Base X

	1994	1995	1996	1997	1998	1999
Officers:	0	525	0	0	0	0
Enlisted:	0	3,282	0	0	0	0
Civilians:	0	199	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	1,000	0	0	0	0
Mil Light Vehic:	0	335	0	0	0	0
Heavy/Spec Vehic:	0	626	0	0	0	0

Transfers from Base X to Shaw, SC

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

Transfers from Barksdale, LA to Eglin, FL

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

Transfers from Eglin, FL to Barksdale, LA

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

INPUT SCREEN THREE - MOVEMENT TABLE (COBRA v4.04) - Page 10
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Transfers from Barksdale, LA to Base X

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

Transfers from Base X to Barksdale, LA

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

INPUT SCREEN THREE - MOVEMENT TABLE (COBRA v4.04) - Page 11
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Transfers from Eglin, FL to Base X

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

Transfers from Base X to Eglin, FL

	1994	1995	1996	1997	1998	1999
Officers:	0	0	0	0	0	0
Enlisted:	0	0	0	0	0	0
Civilians:	0	0	0	0	0	0
Students:	0	0	0	0	0	0
Missn Eqpt (tons):	0	0	0	0	0	0
Suppt Eqpt (tons):	0	0	0	0	0	0
Mil Light Vehic:	0	0	0	0	0	0
Heavy/Spec Vehic:	0	0	0	0	0	0

INPUT SCREEN FOUR - STATIC BASE INFO (COBRA v4.04) - Page 12
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Name: Tyndall, FL

Homeowner Assistance Program: Yes
Unique Activity Information: No

Total Officer Employees:	754
Total Enlisted Employees:	3,722
Total Student Employees:	0
Percent of Military Families Living On Base:	32.0% Total
Civilian Employees:	964
Percent of Civilians Not Willing To Move:	10.0% Officer
Housing Units Available:	0
Enlisted Housing Units Available:	0
Total Base Facilities (Square Feet):	3,204,964
Total Acreage on Base (Acres):	28,824
Officer Variable Housing Allowance ($/Month):	66
Enlisted Variable Housing Allowance ($/Month):	21
Per Diem Rate ($/Day):	78
Freight Cost ($/Ton/Mile):	0.10
Area Cost Factor:	0.85
RPMA Non-Payroll Costs ($K/Year):	14,300
RPMA Payroll Costs ($K/Year):	0
Communications Costs ($K/Year):	1,500 Base
Ops Non-Payroll Costs ($K/Year):	9,100
Base Ops Payroll Costs ($K/Year):	0
Family Housing Costs ($K/Year):	4,500
CHAMPUS On-Base In-Patient Cost/Visit ($):	200
CHAMPUS On-Base Out-Patient Cost/Visit ($):	20
CHAMPUS Shift To Medicare	0.0%

Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Name: Shaw, SC

Homeowner Assistance Program: Yes
Unique Activity Information: No

Total Officer Employees:	705
Total Enlisted Employees:	4,465
Total Student Employees:	0
Percent of Military Families Living On Base:	30.0%
Total Civilian Employees:	556
Percent of Civilians Not Willing To Move:	10.0%
Officer Housing Units Available:	0
Enlisted Housing Units Available:	0
Total Base Facilities (Square Feet):	2,643,777
Total Acreage on Base (Acres):	3,306
Officer Variable Housing Allowance ($/Month):	0
Enlisted Variable Housing Allowance ($/Month):	20
Per Diem Rate ($/Day):	79
Freight Cost ($/Ton/Mile):	0.10
Area Cost Factor:	0.83
RPMA Non-Payroll Costs ($K/Year):	18,700
RPMA Payroll Costs ($K/Year):	0
Communications Costs ($K/Year):	2,500
Base Ops Non-Payroll Costs ($K/Year):	7,100
Base Ops Payroll Costs ($K/Year):	0
Family Housing Costs ($K/Year):	8,900
CHAMPUS On-Base In-Patient Cost/Visit ($):	1,000
CHAMPUS On-Base Out-Patient Cost/Visit ($):	100
CHAMPUS Shift To Medicare	0.0%

INPUT SCREEN FOUR - STATIC BASE INFO (COBRA v4.04) - Page 14
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Name: Barksdale, LA

Homeowner Assistance Program: Yes
Unique Activity Information: No

Total Officer Employees:	822
Total Enlisted Employees:	4,559
Total Student Employees:	0
Percent of Military Families Living On Base:	12.2%
Total Civilian Employees:	1,265
Percent of Civilians Not Willing To Move:	10.0%
Officer Housing Units Available:	0
Enlisted Housing Units Available:	0
Total Base Facilities (Square Feet):	3,682,672
Total Acreage on Base (Acres):	22,361
Officer Variable Housing Allowance ($/Month):	52
Enlisted Variable Housing Allowance ($/Month):	24
Per Diem Rate ($/Day):	83
Freight Cost ($/Ton/Mile):	0.10
Area Cost Factor:	0.86
RPMA Non-Payroll Costs ($K/Year):	15,200
RPMA Payroll Costs ($K/Year):	0
Communications Costs ($K/Year):	300
Base Ops Non-Payroll Costs ($K/Year):	6,600
Base Ops Payroll Costs ($K/Year):	0
Family Housing Costs ($K/Year):	5,200
CHAMPUS On-Base In-Patient Cost/Visit ($):	1,000
CHAMPUS On-Base Out-Patient Cost/Visit ($):	100
CHAMPUS Shift To Medicare	0.0%

INPUT SCREEN FOUR - STATIC BASE INFO (COBRA v4.04) - Page 15
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Name: Eglin, FL

Homeowner Assistance Program: Yes
Unique Activity Information: No

Total Officer Employees:	1,561
Total Enlisted Employees:	6,867
Total Student Employees:	0
Percent of Military Families Living On Base:	20.0%
Total Civilian Employees:	4,050
Percent of Civilians Not Willing To Move:	10.0%
Officer Housing Units Available:	0
Enlisted Housing Units Available:	0
Total Base Facilities (Square Feet):	7,270,818
Total Acreage on Base (Acres):	455,948
Officer Variable Housing Allowance ($/Month):	55
Enlisted Variable Housing Allowance ($/Month):	39
Per Diem Rate ($/Day):	88
Freight Cost ($/Ton/Mile):	0.10
Area Cost Factor:	0.83
RPMA Non-Payroll Costs ($K/Year):	28,000
RPMA Payroll Costs ($K/Year):	0
Communications Costs ($K/Year):	400
Base Ops Non-Payroll Costs ($K/Year):	20,100
Base Ops Payroll Costs ($K/Year):	0
Family Housing Costs ($K/Year):	12,000
CHAMPUS On-Base In-Patient Cost/Visit ($):	0
CHAMPUS On-Base Out-Patient Cost/Visit ($):	0
CHAMPUS Shift To Medicare	0.0%

INPUT SCREEN FOUR - STATIC BASE INFO (COBRA v4.04) - Page 16
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Name: Base X

Homeowner Assistance Program: Yes
Unique Activity Information: No

Total Officer Employees:	564
Total Enlisted Employees:	3,892
Total Student Employees:	0
Percent of Military Families Living On Base:	32.0%
Total Civilian Employees:	3,892
Percent of Civilians Not Willing To Move:	10.0%
Officer Housing Units Available:	0
Enlisted Housing Units Available:	0
Total Base Facilities (Square Feet):	3,023,266
Total Acreage on Base (Acres):	12,006
Officer Variable Housing Allowance ($/Month):	71
Enlisted Variable Housing Allowance ($/Month):	41
Per Diem Rate ($/Day):	77
Freight Cost ($/Ton/Mile):	0.10
Area Cost Factor:	1.00
RPMA Non-Payroll Costs ($K/Year):	17,205
RPMA Payroll Costs ($K/Year):	0
Communications Costs ($K/Year):	1,897
Base Ops Non-Payroll Costs ($K/Year):	5,936
Base Ops Payroll Costs ($K/Year):	0
Family Housing Costs ($K/Year):	4,124
CHAMPUS On-Base In-Patient Cost/Visit ($):	1,000
CHAMPUS On-Base Out-Patient Cost/Visit ($):	100
CHAMPUS Shift To Medicare	20.9%

INPUT SCREEN FIVE - DYNAMIC BASE INFO (COBRA v4.04) - Page 17
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Name: Tyndall, FL

	1994	1995	1996	1997	1998	1999	
1-Time Unique($K):	0	0	0	0	0	0	
1-Time Moving($K):	0	0	0	0	0	0	
Env Mitig Req($K):	0	0	0	0	0	0	
Act Misn Cost($K):	0	0	0	0	0	0	
Misc Rec Cost($K):	0	0	0	0	0	0	
Property (Acres):	0	0	0	0	0	0	
Property ($K):	0	0	-2,695	-2,695	-2,695	-2,695	
Construc Sched(%):	100%	0%	0%	0%	0%	0%	
Shutdown Sched(%):	0%	50%	50%	0%	0%	0%	
Constr Avoid ($K):	2,428	2,260	1,315	2,080	6,831	0	
FamHousAvoid ($K):		0	0	0	0	0	0
Procur Avoid ($K):		0	0	0	0	0	0

(Positive indicates buys, negative indicates sales)

Facility Shut Down (SqFt): 3,204,964
Percent of Family Housing ShutDown: 100.0%

Name: Shaw, SC

	1994	1995	1996	1997	1998	1999
1-Time Unique($K):	0	0	0	0	0	0
1-Time Moving($K):	0	0	0	0	0	0
Env Mitig Req($K):	0	0	0	0	0	0
Act Misn Cost($K):	0	0	0	0	0	0
Misc Rec Cost($K):	0	0	0	0	0	0
Property (Acres):	0	0	0	0	0	0
Property ($K):	0	0	0	0	0	0
Construc Sched(%):	100%	0%	0%	0%	0%	0%
Shutdown Sched(%):	0%	100%	0%	0%	0%	0%
Constr Avoid ($K):	0	0	0	0	0	0
FamHousAvoid ($K):	0	0	0	0	0	0
Procur Avoid ($K):	0	0	0	0	0	0

(Positive indicates buys, negative indicates sales)

Facility Shut Down (SqFt): 0
Percent of Family Housing ShutDown: 0.0%

INPUT SCREEN FIVE - DYNAMIC BASE INFO (COBRA v4.04) - Page 18
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Name: Barksdale, LA

	1994	1995	1996	1997	1998	1999
1-Time Unique($K):	0	0	0	0	0	0
1-Time Moving($K):	0	0	0	0	0	0
Env Mitig Req($K):	0	0	0	0	0	0
Act Misn Cost($K):	0	0	0	0	0	0
Misc Rec Cost($K):	0	0	0	0	0	0
Property (Acres):	0	0	0	0	0	0
Property ($K):	0	0	0	0	0	0

(Positive indicates buys, negative indicates sales)

	1994	1995	1996	1997	1998	1999
Construc Sched(%):	100%	0%	0%	0%	0%	0%
Shutdown Sched(%):	0%	100%	0%	0%	0%	0%
Constr Avoid ($K):	0	0	0	0	0	0
FamHousAvoid ($K):	0	0	0	0	0	0
Procur Avoid ($K):	0	0	0	0	0	0

Facility Shut Down (SqFt): 0
Percent of Family Housing ShutDown: 0.0%

Name: Eglin, FL

	1994	1995	1996	1997	1998	1999
1-Time Unique($K):	0	0	0	0	0	0
1-Time Moving($K):	0	0	0	0	0	0
Env Mitig Req($K):	0	0	0	0	0	0
Act Misn Cost($K):	0	0	0	0	0	0
Misc Rec Cost($K):	0	0	0	0	0	0
Property (Acres):	0	0	0	0	0	0
Property ($K):	0	0	0	0	0	0

(Positive indicates buys, negative indicates sales)

	1994	1995	1996	1997	1998	1999
Construc Sched(%):	100%	0%	0%	0%	0%	0%
Shutdown Sched(%):	0%	100%	0%	0%	0%	0%
Constr Avoid ($K):	0	0	0	0	0	0
FamHousAvoid ($K):	0	0	0	0	0	0
Procur Avoid ($K):	0	0	0	0	0	0

Facility Shut Down (SqFt): 0
Percent of Family Housing ShutDown: 0.0%

INPUT SCREEN FIVE - DYNAMIC BASE INFO (COBRA v4.04) - Page 19
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Name: Base X

	1994	1995	1996	1997	1998	1999
1-Time Unique($K):	0	0	0	0	0	0
1-Time Moving($K):	0	0	0	0	0	0
Env Mitig Req($K):	0	0	0	0	0	0
Act Misn Cost($K):	0	0	0	0	0	0
Misc Rec Cost($K):	0	0	0	0	0	0
Property (Acres):	0	0	0	0	0	0
Property ($K):	0	0	0	0	0	0

(Positive indicates buys, negative indicates sales)

	1994	1995	1996	1997	1998	1999
Construc Sched(%):	100%	0%	0%	0%	0%	0%
Shutdown Sched(%):	0%	100%	0%	0%	0%	0%
Constr Avoid ($K):	0	0	0	0	0	0
FamHousAvoid ($K):	0	0	0	0	0	0
Procur Avoid ($K):	0	0	0	0	0	0

Facility Shut Down (SqFt): 0
Percent of Family Housing ShutDown: 0.0%

Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Name: Tyndall, FL

	1994	1995	1996	1997	1998	1999
Officer FS Chg:	0	0	0	0	0	0
Enlisted FS Chg:	0	0	0	0	0	0
Civilian FS Chg:	0	0	0	0	0	0
Officers Elim:	0	53	0	0	0	0
Enlisted Elim:	0	793	0	0	0	0
Civilians Elim:	0	374	0	0	0	0
Caretakers - Mil:	0	0	0	0	0	0
Caretakers - Civ:	0	0	0	0	0	0
CHAMPUS InPat/Yr:	0	0	-8,400	-8,400	-8,400	-8,400
CHAMPUS OutPat/Yr:	0	0	-84	-84	-84	-84

Name: Shaw, SC

	1994	1995	1996	1997	1998	1999
Officer FS Chg:	0	0	0	0	0	0
Enlisted FS Chg:	0	0	0	0	0	0
Civilian FS Chg:	0	0	0	0	0	0
Officers Elim:	0	0	0	0	0	0
Enlisted Elim:	0	0	0	0	0	0
Civilians Elim:	0	0	0	0	0	0
Caretakers - Mil:	0	0	0	0	0	0
Caretakers - Civ:	0	0	0	0	0	0
CHAMPUS InPat/Yr:	0	0	0	0	0	0
CHAMPUS OutPat/Yr:	0	0	0	0	0	0

Name: Barksdale, LA

	1994	1995	1996	1997	1998	1999
Officer FS Chg:	0	0	0	0	0	0
Enlisted FS Chg:	0	0	0	0	0	0
Civilian FS Chg:	0	0	0	0	0	0
Officers Elim:	0	0	0	0	0	0
Enlisted Elim:	0	0	0	0	0	0
Civilians Elim:	0	0	0	0	0	0
Caretakers - Mil:	0	0	0	0	0	0
Caretakers - Civ:	0	0	0	0	0	0
CHAMPUS InPat/Yr:	0	0	0	0	0	0
CHAMPUS OutPat/Yr:	0	0	0	0	0	0

INPUT SCREEN SIX - BASE PERSONNEL INFO (COBRA v4.04) - Page 21
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Name: Eglin, FL

	1994	1995	1996	1997	1998	1999
Officer FS Chg:	0	0	0	0	0	0
Enlisted FS Chg:	0	0	0	0	0	0
Civilian FS Chg:	0	0	0	0	0	0
Officers Elim:	0	0	0	0	0	0
Enlisted Elim:	0	0	0	0	0	0
Civilians Elim:	0	0	0	0	0	0
Caretakers - Mil:	0	0	0	0	0	0
Caretakers - Civ:	0	0	0	0	0	0
CHAMPUS InPat/Yr:	0	0	0	0	0	0
CHAMPUS OutPat/Yr:	0	0	0	0	0	0

Name: Base X

	1994	1995	1996	1997	1998	1999
Officer FS Chg:	0	0	0	0	0	0
Enlisted FS Chg:	0	0	0	0	0	0
Civilian FS Chg:	0	0	0	0	0	0
Officers Elim:	0	0	0	0	0	0
Enlisted Elim:	0	0	0	0	0	0
Civilians Elim:	0	0	0	0	0	0
Caretakers - Mil:	0	0	0	0	0	0
Caretakers - Civ:	0	0	0	0	0	0
CHAMPUS InPat/Yr:	0	0	0	0	0	0
CHAMPUS OutPat/Yr:	0	0	0	0	0	0

INPUT SCREEN SEVEN - MILCON BASE INFO (COBRA v4.04) - Page 22
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Name: Tyndall, FL

Description	Category	New Con	Rehab	Cost ($K)
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0

Name: Shaw, SC

Description	Category	New Con	Rehab	Cost ($K)
Clearing Shaw	OptCat-A	0	0	66,900
Milcon	OptCat-A	0	0	38,300
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0

Name: Barksdale, LA

Description	Category	New Con	Rehab	Cost ($K)
Milcon	OptCat-A	0	0	45,500
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0

INPUT SCREEN SEVEN - MILCON BASE INFO (COBRA v4.04) - Page 24
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Name: Eglin, FL

Description	Category	New Con	Rehab	Cost ($K)
Milcon	OptCat-A	0	0	66,100
Milcon	OptCat-A	0	0	31,800
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0

Name: Base X

Description	Category	New Con	Rehab	Cost ($K)
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0
	OptCat-A	0	0	0

STANDARD PERSONNEL FACTORS (COBRA v4.04) - Page 25
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Percentage of Officers Married	71.00%
Percentage of Enlisted Married	54.00%
Enlisted Housing Military Construction	98.50%
Officer Salary ($/Year)	75,421.00
Officer BAQ with Dependents	7,877.00
Enlisted Salary ($/Year)	34,875.00
Enlisted BAQ with Dependents	5,313.00
Average Unemployment Cost ($/Week)	200.00
Unemployment Eligibility Period (Weeks)	16
Civilian Salary ($)	44,221.00
Civilian Turnover Rate	15.00%
Civilian Early Retirement Rate	10.00%
Civilian Quitting Rate	3.00%
Civilian RIF Pay Factor	11.22%
Civilian Retirement Pay Factor	8.80%
Priority Placement Service	30.00%
PPS Actions Involving PCS	40.00%
Civilian PCS Costs ($)	25,789.00
New Hire Cost ($)	0.00
National Median Home Price ($)	109,010.00
Home Sale Reimbursement	10.00%
Maximum Home Sale Reimbursement ($)	22,193.00
Home Purch Reimbursement	5.00%
Maximum Home Purch Reimbursement ($)	11,096.00
Civilian Homeowning Rate	64.00%
HAP Home Value Reimbursement Rate	37.50%
HAP Homeowner Receiving Rate	20.00%
RSE Home Value Reimbursement Rate	23.00%
RSE Homeowner Receiving Rate	12.00%

Standard Factors File Description: **BRAC 93 FACTORS**

STANDARD FACILITY FACTORS (COBRA v4.04) - Page 26
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

RPMA Building SF Cost Index	0.64
BOS Index (RPMA vs population)	0.56
(Indices are used as exponents)	
Support for Move Factor	0.10%

Caretaker Costs:
--------- ------

Administrative Space Needs (SF/Caretaker)	162.00
Percentage of Original RPMA Cost	23.60%
Mothball Cost ($/SqFt)	0.00
Discount Rate for NPV.RPT/ROI:	7.0%
Inflation Rate for NPV.RPT/ROI:	0.0%

Inflation Rate	1994	1995	1996	1997	1998	1999 for
FINANCE.RPT:	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Average Bachelor Quarters Size (SF):	250.00
Average Family Quarters Size (SF):	1,819.00
Rehabilitation Cost vs. New Construction Cost	75.00%
Information Management Account	15.00%
Design Rate	9.00%
Supervision, Inspection, OverHead Rate	6.00%
Contingency Planning Rate	7.00%
Site Preparation Rate	23.00%

STANDARD TRANSPORTATION FACTORS (COBRA v4.04) - Page 27
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Material per Assigned Person (Lbs)		710
HHG Weight Per Officer Family (Lb)		15,284.00
HHG Weight Per Enlisted Family (Lb)		9,230.00
HHG Weight Per Military Single (Lb)		4,927.00
HHG Weight Per Civilian (Lb)	18,000.00	
Household Goods Cost ($/100Lb)		31.10
(Includes Packing, Unpacking, Storage, and Misc. Costs)		
Shipping Loss Rate		7.4%
Equipment Packing & Crating Cost ($/Ton)		344.00
Military Light Vehicle Cost ($/Mile)		0.09
Heavy or Special Vehicle Cost ($/Mile)		0.09
Pers Owned Vehic Reimburse ($/Mile)		0.18
Air Transport Per Passenger Mile ($)		0.18
Misc Expenses Per Direct Employee ($)		700.00
Avg Military Service Tour Length (Years)		2.20
Routine PCS Costs/Person/Tour ($)		4,231.00
One-Time Officer PCS Cost ($)		0.00
One-Time Enlisted PCS Cost ($)		0.00

STANDARD CONSTRUCTION FACTORS (COBRA v4.04) - Page 28
Data As Of 18:11 01/30/1993, Report Created 09:26 06/25/1993

Category:	Units:	Cost/UM($):
Horizontal	(SY)	37
Waterfront	(LF)	91
Air Operations	(SF)	125
Operational	(SF)	115
Administrative	(SF)	102
School Buildings	(SF)	105
Maintenance Shops	(SF)	104
Bachelor Quarters	(EA)	19,510
Family Quarters	(EA)	115,052
Covered Storage	(SF)	58
Dining Facilities	(SF)	174
Recreation Facilities	(SF)	123
Communications Facilities	(SF)	0
Shipyard Maintenance	(SF)	0
RDT & E Facilities	(SF)	0
POL Storage	(BL)	0
Ammunition Storage	(SF)	0
Medical Facilities	(SF)	0
Environmental	( )	0
Optional Category A	( )	0
Optional Category B	( )	0
Optional Category C	( )	0
Optional Category D	( )	0
Optional Category E	( )	0
Optional Category F	( )	0
Optional Category G	( )	0
Optional Category H	( )	0
Optional Category I	( )	0
Optional Category J	( )	0
Optional Category K	( )	0
Optional Category L	( )	0
Optional Category M	( )	0
Optional Category N	( )	0
Optional Category O	( )	0
Optional Category P	( )	0

Close Base FY 95/4: Realign to bases specified

Distance=1,000 mi, @ $1.00 ton/mile

Base X & Base Y values average of small aircraft bases

Appendix D: Preliminary Data Set
(All Data -- Air Force, AFRES, ANG, Depot, Army)

CLOSURE BASE	*NPV*	*Miles (pers)*	*Equip(T)*	*Vehicles*	*Officers*	*Enlisted*
Whiteman	($85,447,000)	476	3,000	787	357	2,661
Tyndall	($337,166,000)	452	3,000	550	754	3,722
Travis	$117,950,000	304	3,000	550	1,032	5,549
Shaw_II	($256,303,000)	1,213	3,000	823	761	5,184
Shaw_I	($456,926,000)	851	2,999	960	705	4,465
Seymour_Johnson	($280,000,000)	1,491	3,000	961	534	3,925
Pope	($314,859,000)	209	3,000	685	422	3,697
Plattsburgh	($470,858,000)	2,114	3,000	626	302	1,793
Mountain_Home	($224,988,000)	671	3,000	685	424	2,966
Moody_II	($248,356,000)	1,704	2,997	461	338	3,015
Moody_I	($453,257,000)	1,148	3,000	525	265	2,145
Minot	($511,371,000)	901	2,997	342	680	3,847
McGuire	($544,376,000)	404	3,000	330	332	2,957
McConnell	($295,355,000)	657	3,001	880	343	2,357
McChord	($133,457,000)	612	3,000	622	414	3,555
March	($247,492,000)	639	3,000	592	546	2,681
Malmstrom	$56,950,000	1,250	3,000	672	529	2,780
Luke	($34,325,000)	695	3,000	486	641	5,126
Little_Rock	($222,361,000)	1,061	3,750	813	595	3,360
K_I_Sawyer	($348,897,000)	1,505	3,000	695	302	2,052
Homestead	($940,268,000)	1,163	3,000	624	418	3,447
Holloman	($353,603,000)	770	3,000	483	554	3,993
Griffiss	($295,612,000)	812	2,997	803	595	3,165
Grand_Forks	($610,729,000)	675	3,000	1,008	680	4,030
Fairchild	($113,853,000)	1,473	3,000	899	709	3,923
Ellsworth	($324,505,000)	806	3,000	991	633	3,647
Dyess	$87,703,000	861	0	516	752	4,328
Dover	($140,220,000)	527	3,000	726	375	3,608
Davis_Monthan	($360,473,000)	853	2,820	0	746	4,145
Charleston	($212,911,000)	1,029	3,000	642	576	3,389
Cannon	($188,682,000)	1,918	3,000	610	543	4,501
Beale	($347,585,000)	1,556	3,000	616	348	2,295
Barksdale	($151,376,000)	1,114	3,110	542	836	4,559
Altus	($269,672,000)	1,668	3,000	590	372	2,495

CLOSURE BASE	Civilians	Facilities (SF)	Acreage	Area Cost Factor	RPMA
Whiteman	611	2,736,279	3,542	1.11	$10,400,000
Tyndall	964	3,204,964	28,824	0.85	$14,300,000
Travis	2,124	5,440,672	6,272	1.24	$18,400,000
Shaw_II	556	2,643,777	3,306	0.83	$13,600,000
Shaw_I	556	2,643,777	3,306	0.83	$13,600,000
Seymour_Johnson	553	2,508,497	3,233	0.80	$10,200,000
Pope	365	1,673,384	1,847	0.80	$6,300,000
Plattsburgh	352	2,474,133	3,440	1.11	$11,000,000
Mountain_Home	505	2,276,607	6,721	1.10	$8,100,000
Moody_II	445	1,646,829	5,039	0.77	$8,200,000
Moody_I	445	1,646,829	5,039	0.77	$8,200,000
Minot	568	3,038,678	4,714	1.07	$11,900,000
McGuire	1,876	4,043,220	3,597	1.15	$17,800,000
McConnell	881	2,584,320	2,594	0.92	$9,500,000
McChord	1,158	3,504,024	4,616	1.00	$7,300,000
March	1,651	3,393,867	6,594	1.26	$12,100,000
Malmstrom	321	2,523,207	3,608	1.20	$13,800,000
Luke	1,085	3,221,515	4,198	0.96	$15,500,000
Little_Rock	644	2,660,925	6,128	0.79	$7,800,000
K_I_Sawyer	351	2,126,652	5,214	1.21	$8,800,000
Homestead	912	2,272,571	3,345	0.90	$13,800,000
Holloman	989	4,309,737	50,999	1.01	$14,300,000
Griffiss	2,320	4,356,692	3,896	1.05	$13,600,000
Grand_Forks	481	2,872,831	5,418	0.96	$15,500,000
Fairchild	576	4,167,580	5,000	0.99	$9,400,000
Ellsworth	497	4,193,219	6,217	1.02	$15,500,000
Dyess	416	2,757,251	6,407	0.92	$11,100,000
Dover	1,257	3,214,109	3,900	1.02	$12,400,000
Davis_Monthan	1,408	3,569,390	10,613	0.90	$16,400,000
Charleston	1,092	2,875,783	6,164	0.91	$9,300,000
Cannon	498	2,115,974	3,781	1.10	$7,700,000
Beale	400	2,205,401	22,944	0.96	$7,600,000
Barksdale	2,196	3,682,672	22,361	0.86	$15,200,000
Altus	475	2,229,839	2,806	0.86	$7,800,000

CLOSURE BASE	Comm Budget	Base Ops Cost	MFH Budget	CHAMPUS (in-pat)
Whiteman	$1,000,000	$5,900,000	$6,000,000	($2,400,000)
Tyndall	$1,500,000	$9,100,000	$4,500,000	($6,720,000)
Travis	$1,200,000	$8,700,000	$5,200,000	($21,600,000)
Shaw_II	$2,500,000	$7,100,000	$6,000,000	($5,520,000)
Shaw_I	$2,500,000	$7,100,000	$6,000,000	($5,520,000)
Seymour_Johnson	$1,300,000	$5,300,000	$5,500,000	($3,520,000)
Pope	$1,300,000	$7,200,000	$1,600,000	$0
Plattsburgh	$500,000	$6,100,000	$4,900,000	($800,000)
Mountain_Home	$1,600,000	$6,800,000	$4,800,000	($4,000,000)
Moody_II	$1,300,000	$3,700,000	$1,300,000	($4,000,000)
Moody_I	$1,300,000	$3,700,000	$1,300,000	($4,000,000)
Minot	$700,000	$6,100,000	$8,300,000	($2,480,000)
McGuire	$2,000,000	$12,000,000	$6,500,000	($6,400,000)
McConnell	$600,000	$6,400,000	$2,500,000	($2,000,000)
McChord	$800,000	$13,230,000	$3,000,000	$0
March	$500,000	$5,800,000	$5,300,000	($8,800,000)
Malmstrom	$700,000	$4,900,000	$7,300,000	$0
Luke	$2,100,000	$6,800,000	$4,300,000	($9,600,000)
Little_Rock	$1,600,000	$5,400,000	$5,900,000	($4,320,000)
K_I_Sawyer	$1,000,000	$5,100,000	$4,300,000	($1,440,000)
Homestead	$2,600,000	$7,100,000	$5,500,000	($4,800,000)
Holloman	$2,200,000	$9,400,000	$5,400,000	($2,400,000)
Griffiss	$400,000	$5,900,000	$4,900,000	($4,800,000)
Grand_Forks	$700,000	$6,400,000	$7,400,000	$0
Fairchild	$1,000,000	$6,600,000	$5,000,000	($5,600,000)
Ellsworth	$700,000	$6,400,000	$7,400,000	($1,600,000)
Dyess	$700,000	$5,300,000	$3,500,000	($3,200,000)
Dover	$1,000,000	$6,300,000	$5,400,000	($4,800,000)
Davis_Monthan	$1,900,000	$6,300,000	$3,400,000	($8,800,000)
Charleston	$1,200,000	$6,900,000	$2,900,000	$0
Cannon	$1,600,000	$5,800,000	$3,100,000	$0
Beale	$900,000	$5,000,000	$6,900,000	($2,400,000)
Barksdale	$300,000	$6,600,000	$5,200,000	($8,800,000)
Altus	$600,000	$4,300,000	$2,300,000	($1,600,000)

CLOSURE BASE	CHAMPUS (out-pat)	One Time Move	Prop. Trans	Net Constr.
Whiteman	($239,600)	$0	$0	$119,930,000
Tyndall	($6,720)	$0	($10,780,000)	$233,686,000
Travis	($713,680)	$1,750,000	($8,358,000)	$731,025,000
Shaw_II	($77,280)	$2,105,000	$0	$295,698,000
Shaw_I	($77,280)	$2,105,000	$0	$125,798,000
Seymour_Johnson	($34,560)	$2,820,000	$0	$257,157,000
Pope	$0	$0	$0	$149,796,000
Plattsburgh	($797,120)	$250,000	$0	$65,661,000
Mountain_Home	($361,280)	$0	$0	$290,706,000
Moody_II	($61,680)	$0	$0	$138,167,000
Moody_I	($61,680)	$0	$0	($41,533,000)
Minot	($10,000)	$4,433,000	$0	$83,677,000
McGuire	($352,720)	$1,500,000	$0	$195,963,000
McConnell	($49,760)	$2,325,000	$0	$200,371,000
McChord	$0	$2,650,000	$0	$355,337,000
March	($421,200)	$2,191,000	$0	$177,618,000
Malmstrom	$0	$77,000,000	$0	$230,611,000
Luke	($687,520)	$9,700,000	$0	$391,876,000
Little_Rock	($68,800)	$8,500,000	($7,113,000)	$345,492,000
K_I_Sawyer	($14,960)	$1,130,000	$0	$100,951,000
Homestead	($536,000)	$25,000	$0	($363,300,000)
Holloman	($13,920)	$610,000	$0	$357,331,000
Griffiss	($92,480)	$21,100,000	$0	$292,040,000
Grand_Forks	($3,853,200)	$3,500,000	$0	$42,157,000
Fairchild	($488,000)	$1,880,000	$0	$288,439,000
Ellsworth	($389,120)	$4,075,000	$0	$410,231,000
Dyess	($735,600)	$10,230,000	($2,000,000)	$535,170,000
Dover	($34,960)	$2,492,000	($371,100)	$311,090,000
Davis_Monthan	($257,120)	$0	$0	$79,307,000
Charleston	$0	$8,500,000	($5,280,000)	$163,536,000
Cannon	$0	$500,000	$0	$270,854,000
Beale	($450,320)	$7,455,000	$0	$163,171,000
Barksdale	($84,400)	$6,790,000	($11,932,000)	$425,110,000
Altus	($229,600)	$1,950,000	$0	$170,088,000

CLOSURE BASE	NPV	Miles (pers)	Equip(T)	Vehicles	Officers	Enlisted
Bergstrom	($196,499,000)	1,000	3,000	0	112	450
Carswell	($205,872,000)	1,000	3,000	199	171	685
Dobbins	($116,153,000)	1,000	3,000	152	189	756
General_Mitchell	($132,080,000)	1,000	3,000	109	170	681
Grissom	($253,222,000)	1,000	3,000	433	181	725
Greater_Pittsburgh	($120,133,000)	1,000	3,000	103	167	757
Minn_StPaul	($97,372,000)	1,000	3,000	121	157	627
OHare	($172,149,000)	1,000	3,000	135	234	937
Westover	($69,685,000)	1,000	3,000	370	274	1,098
Youngstown	($122,098,000)	1,000	3,000	137	137	548
Boise	($16,075,000)	1,000	3,000	118	193	1,206
Buckley	($117,467,000)	1,000	3,000	239	209	1,272
Ellington	($34,205,000)	1,000	3,000	134	119	746
Fresno	($61,211,000)	1,000	3,000	96	110	913
Great_Falls	($6,556,000)	1,000	3,000	114	104	911
Martin_State	($2,777,000)	1,000	3,000	126	253	1,697
McEntire	($13,054,000)	1,000	3,000	228	135	1,227
Otis	($201,678,000)	1,000	3,000	225	130	1,027
Pittsburgh	$5,481,000	1,000	3,000	118	193	1,206
Portland	($70,374,000)	1,000	3,000	272	146	1,286
Salt_Lake_City	($313,000)	1,000	3,000	143	162	1,439
Selfridge	($260,373,000)	1,000	3,000	334	242	1,640
Stewart	($7,418,000)	1,000	3,000	198	152	1,630
Tucson	$4,171,000	1,000	3,000	157	133	1,286
Hill	$424,725,000	1,294	3,000	0	828	3,558
Kelly	$610,033,000	722	3,000	0	417	3,771
McClellan	($25,363,000)	1,491	3,000	0	33	2,411
Newark	$126,802,000	1,603	0	0	33	64
Robins	$315,695,000	1,639	0	0	725	3,025
Tinker	$278,587,000	806	3,000	1,963	1,413	5,576
Carson	($467,323,000)	1,340	11,337	102,697	1,723	15,667
Drum4	($903,670,000)	1,340	1,500	15,000	1,120	9,137
Drum5	($823,741,000)	1,340	1,500	15,000	1,028	8,438

CLOSURE BASE	Civilians	Facilities (SF)	Acreage	Area Cost Factor	RPMA
Bergstrom	625	516,902	710	1.00	$8,016,000
Carswell	583	891,671	321	1.00	$9,694,000
Dobbins	577	823,691	1,666	1.00	$5,024,000
General_Mitchell	329	304,920	0	1.00	$4,368,000
Grissom	958	1,023,324	2,722	1.00	$7,391,000
Greater_Pittsburgh	341	428,690	176	1.00	$4,209,000
Minn_StPaul	323	1,001,173	266	1.00	$4,000,000
OHare	388	823,691	274	1.00	$7,701,000
Westover	965	1,118,761	2,609	1.00	$25,000
Youngstown	354	402,154	230	1.00	$6,091,000
Boise	360	340,858	576	1.00	$2,932,000
Buckley	477	1,048,820	3,832	1.00	$9,034,000
Ellington	289	463,496	216	1.00	$5,007,000
Fresno	294	262,100	0	1.00	$2,543,000
Great_Falls	290	282,775	0	1.00	$2,580,000
Martin_State	354	317,660	0	1.00	$1,975,000
McEntire	342	400,291	2,472	1.00	$1,688,000
Otis	574	979,462	3,886	1.00	$5,963,000
Pittsburgh	360	310,670	176	1.00	$1,952,000
Portland	352	648,933	289	1.00	$3,503,000
Salt_Lake_City	347	364,079	152	1.00	$2,529,000
Selfridge	990	1,807,764	2,559	1.00	$16,578,000
Stewart	432	779,544	248	1.00	$3,085,000
Tucson	626	393,602	81	1.00	$2,929,000
Hill	9,045	11,537,622	6,646	0.98	$23,300,000
Kelly	14,463	14,862,977	3,996	0.88	$47,700,000
McClellan	8,325	11,125,536	3,856	1.14	$25,700,000
Newark	1,691	743,951	67	0.99	$3,800,000
Robins	11,313	11,795,410	8,720	0.77	$21,900,000
Tinker	11,476	13,250,953	3,905	0.87	$26,900,000
Carson	2,400	11,003,000	373,287	1.03	$25,779,000
Drum4	2,263	11,911,000	107,296	1.18	$23,638,000
Drum5	2,533	11,911,000	107,296	1.18	$23,638,000

CLOSURE BASE	Comm Budget	Base Ops Cost	MFH Budget	CHAMPUS (in-pat)
Bergstrom	$1,441,000	$4,415,000	$0	$0
Carswell	$1,469,000	$5,309,000	$0	$0
Dobbins	$870,000	$2,465,000	$0	$0
General_Mitchell	$2,293,000	$5,672,000	$0	$0
Grissom	$5,793,000	$4,671,000	$0	$0
Greater_Pittsburgh	$1,061,000	$5,389,000	$0	$0
Minn_StPaul	$1,200,000	$3,100,000	$0	$0
OHare	$371,000	$6,695,000	$0	$0
Westover	$11,000	$11,000	$0	$0
Youngstown	$3,653,000	$3,653,000	$0	$0
Boise	$723,000	$0	$0	$0
Buckley	$1,172,000	$258,000	$0	$0
Ellington	$558,000	$0	$0	$0
Fresno	$641,000	$5,630,000	$0	$0
Great_Falls	$273,000	$40,000	$0	$0
Martin_State	$558,000	$0	$0	$0
McEntire	$538,000	$371,000	$0	$0
Otis	$1,335,000	$0	$0	$0
Pittsburgh	$287,000	$0	$0	$0
Portland	$879,000	$2,042,000	$0	$0
Salt_Lake_City	$384,000	$0	$0	$0
Selfridge	$296,000	$0	$0	$0
Stewart	$987,000	$0	$0	$0
Tucson	$341,000	$0	$0	$0
Hill	$2,800,000	$9,800,000	$4,000,000	$0
Kelly	$1,400,000	$7,000,000	$2,300,000	$0
McClellan	$3,900,000	$7,200,000	$3,500,000	$0
Newark	$200,000	$600,000	$0	$0
Robins	$4,400,000	$6,900,000	$5,600,000	$20,000,000
Tinker	$6,100,000	$9,100,000	$1,800,000	$0
Carson	$2,025,000	$66,186,000	$12,350,000	$0
Drum4	$918,000	$95,755,000	$36,082,000	$0
Drum5	$918,000	$95,485,000	$36,082,000	$0

CLOSURE BASE	CHAMPUS (out-pat)	One Time Move	Prop. Trans	Net Constr.
Bergstrom	$0	$1,000,000	$0	$0
Carswell	$0	$1,000,000	$0	$0
Dobbins	$0	$0	$0	$0
General_Mitchell	$0	$0	$0	$0
Grissom	$0	$0	$0	$0
Greater_Pittsburgh	$0	$0	$0	$0
Minn_StPaul	$0	$0	$0	$0
OHare	$0	$0	$0	$0
Westover	$0	$0	$0	$0
Youngstown	$0	$0	$0	$0
Boise	$0	$270,000	$0	$0
Buckley	$0	$0	$0	($3,900,000)
Ellington	$0	$0	$0	$0
Fresno	$0	$1,500,000	$0	$0
Great_Falls	$0	$0	$0	$0
Martin_State	$0	$0	$0	$0
McEntire	$0	$0	$0	$0
Otis	$0	$250,000	$0	$0
Pittsburgh	$0	$270,000	$0	$0
Portland	$0	$0	$0	$0
Salt_Lake_City	$0	$0	$0	$0
Selfridge	$0	$3,000,000	$0	($2,600,000)
Stewart	$0	$0	$0	$0
Tucson	$0	$2,560,000	$0	$0
Hill	$0	$332,083,000	$0	$443,878,000
Kelly	$0	$232,049,000	$0	$751,589,000
McClellan	$0	$145,729,000	$0	$108,598,000
Newark	$0	$131,091,000	$0	$57,508,000
Robins	$2,168,400	$157,722,000	$0	$340,971,000
Tinker	$0	$215,149,000	$0	$480,718,000
Carson	$0	$0	$0	$396,879,160
Drum4	$0	$0	$0	($15,100,000)
Drum5	$0	$0	$0	$8,351,458

CLOSURE BASE	NPV	Miles (pers)	Equip(T)	Vehicles	Officers	Enlisted
Drum6	($529,542,000)	2,647	1,500	15,000	3,361	23,061
FSTC2	$29,102,000	112	20	0	32	7
FSTC3	$30,241,000	81	20	0	32	7
FSTC4	$32,700,000	137	20	0	32	7
FSTC5	$29,218,000	146	20	0	32	7
ITAC	$15,127,000	564	100	40	81	189
ITAC1	$8,699,000	16	10	0	81	189
ITAC2	$22,275,000	16	20	0	81	189
ITAC3	$13,424,000	16	20	0	81	189
ITAC4	$8,064,000	16	20	0	81	189
ITAC5	($16,494,000)	564	100	40	81	189
ITAC7	($1,090,000)	40	20	0	81	189
Polk1	$97,084,000	1,285	4,451	57,549	1,487	10,760
Polk2	($1,259,144,000)	1,340	1,500	30,000	1,221	9,894
Polk3	($1,404,331,000)	1,340	300	1,000	1,231	9,910
Presidio1	$28,183,000	1,311	0	0	242	445
Presidio2	$53,753,000	1,311	0	0	249	434
Presidio3	$28,132,000	50	0	0	242	445
Presidio4	$10,222,000	35	0	0	138	100
Presidio5	$4,411,000	35	0	0	138	100
Presidio6	$7,577,000	35	0	0	138	100
Presidio7	($669,032,000)	35	0	0	138	100
Presidio8	$28,096,000	50	0	0	242	445
Richardson1	($738,399,000)	351	877	19,361	398	3,607
Richardson2	($140,795,000)	360	877	19,361	398	3,607
Richardson3	($592,765,000)	360	877	19,361	398	3,607
Richardson5	($589,542,000)	360	877	19,361	398	3,607
Richardson6	$4,762,000	393	342	5,995	431	3,988
Riley1	$965,428,000	759	0	0	1,708	14,073
Riley2	$979,394,000	759	0	0	1,708	14,073
Riley3	($1,083,110,000)	1,340	3,000	30,000	1,708	14,073
Riley4	($1,148,821,000)	0	3,000	30,000	1,628	13,680
Riley5	($1,059,209,000)	1,160	3,000	30,000	1,628	13,680
Riley6	$560,491,000	742	3,000	45,000	1,703	13,694
Schofield1	($666,946,000)	2,971	700	1,500	1,301	12,215
Schofield2	($281,164,000)	2,995	4,117	69,366	1,301	12,215

CLOSURE BASE	Civilians	Facilities (SF)	Acreage	Area Cost Factor	RPMA
Drum6	4,532	23,731,000	346,012	1.00	$30,318,000
FSTC2	520	1	0	0.92	$1,202,000
FSTC3	520	1	0	0.92	$1,202,000
FSTC4	520	1	0	0.92	$1,202,000
FSTC5	520	1	0	0.92	$1,202,000
ITAC	336	1	0	1.05	$547,000
ITAC1	336	0	0	1.05	$0
ITAC2	336	1	0	1.05	$547,000
ITAC3	336	1	0	1.05	$547,000
ITAC4	336	1	0	1.05	$547,000
ITAC5	336	1	0	1.05	$547,000
ITAC7	336	1	0	1.00	$0
Polk1	3,017	16,831,000	198,399	0.95	$14,018,000
Polk2	2,867	16,831,000	198,399	0.95	$14,018,000
Polk3	2,865	16,831,000	198,399	0.95	$22,399,000
Presidio1	1,453	0	0	1.00	$0
Presidio2	1,453	0	0	1.00	$0
Presidio3	1,453	0	0	1.00	$0
Presidio4	105	0	1,280	1.39	$0
Presidio5	105	0	1,280	1.39	$0
Presidio6	105	0	1,280	1.39	$0
Presidio7	105	0	1,280	1.39	$0
Presidio8	1,453	0	0	1.00	$0
Richardson1	1,393	7,695,000	61,278	1.69	$10,016,000
Richardson2	1,393	7,695,000	61,278	1.69	$10,016,000
Richardson3	1,393	7,695,000	61,278	1.69	$10,016,000
Richardson5	1,393	7,695,000	61,278	1.69	$10,016,000
Richardson6	1,096	7,695,000	61,278	1.69	$10,016,000
Riley1	2,915	14,105,000	100,667	0.90	$15,022,000
Riley2	2,915	14,105,000	100,667	0.90	$23,478,000
Riley3	2,328	14,105,000	100,667	0.90	$23,478,000
Riley4	2,287	14,105,000	100,667	0.90	$23,478,000
Riley5	2,816	14,105,000	100,667	0.90	$23,478,000
Riley6	2,321	14,105,000	100,667	0.90	$15,022,000
Schofield1	963	12,310,000	101,600	1.42	$20,442,000
Schofield2	963	12,310,000	101,600	1.42	$20,442,000

CLOSURE BASE	Comm Budget	Base Ops Cost	MFH Budget	CHAMPUS (in-pat)
Drum6	$2,017,000	$76,267,000	$20,114,000	$0
FSTC2	$413,000	$0	$0	$0
FSTC3	$413,000	$0	$0	$0
FSTC4	$413,000	$0	$0	$0
FSTC5	$413,000	$0	$0	$0
ITAC	$0	$0	$0	$0
ITAC1	$0	$0	$0	$0
ITAC2	$0	$0	$0	$0
ITAC3	$0	$0	$0	$0
ITAC4	$0	$0	$0	$0
ITAC5	$0	$0	$0	$0
ITAC7	$0	$547,000	$0	$0
Polk1	$1,328,000	$50,907,000	$20,083,000	$0
Polk2	$1,328,000	$50,907,000	$20,083,000	$0
Polk3	$664,000	$57,633,000	$20,083,000	$0
Presidio1	$0	$0	$0	$0
Presidio2	$0	$0	$0	$0
Presidio3	$0	$0	$0	$0
Presidio4	$0	$0	$0	$0
Presidio5	$0	$0	$0	$0
Presidio6	$0	$0	$0	$0
Presidio7	$0	$0	$0	$0
Presidio8	$0	$0	$0	$0
Richardson1	$1,460,000	$47,642,000	$14,587,000	$0
Richardson2	$1,460,000	$47,642,000	$14,587,000	$0
Richardson3	$1,460,000	$47,642,000	$14,587,000	$0
Richardson5	$1,460,000	$47,642,000	$14,587,000	$0
Richardson6	$1,460,000	$47,642,000	$14,587,000	$0
Riley1	$1,239,000	$51,197,000	$20.606,000	$0
Riley2	$870,000	$62,511,000	$20,606,000	$0
Riley3	$870,000	$62,511,000	$20,606,000	$0
Riley4	$870,000	$62,511,000	$20,606,000	$0
Riley5	$870,000	$62,511,000	$20,606,000	$0
Riley6	$1,239,000	$51,197,000	$20,606,000	$0
Schofield1	$292,000	$51,248,000	$16,568,000	$0
Schofield2	$292,000	$51,248,000	$16,568,000	$0

CLOSURE BASE	CHAMPUS (out-pat)	One Time Move	Prop. Trans	Net Constr.
Drum6	$0	$0	$0	$19,075,510
FSTC2	$0	$0	$0	$29,906,000
FSTC3	$0	$0	$0	$32,966,000
FSTC4	$0	$0	$0	$32,966,000
FSTC5	$0	$0	$0	$32,966,000
ITAC	$0	$0	$0	$37,000,000
ITAC1	$0	$0	$0	$0
ITAC2	$0	$0	$0	$21,948,000
ITAC3	$0	$0	$0	$12,898,000
ITAC4	$0	$0	$0	$7,000,000
ITAC5	$0	$0	$0	$37,000,000
ITAC7	$0	$0	$0	$11,958,000
Polk1	$0	$0	$0	$1,398,906,256
Polk2	$0	$0	$0	$111,699,300
Polk3	$0	$0	$0	$0
Presidio1	$0	$0	$0	$23,174,000
Presidio2	$0	$124,000	$0	$23,174,000
Presidio3	$0	$320,000	$0	$6,171,000
Presidio4	$0	$506,000	$0	$6,171,000
Presidio5	$0	$506,000	$0	($17,158,000)
Presidio6	$0	($6,065,000)	$0	($14,058,000)
Presidio7	$0	$506,000	$0	($17,158,000)
Presidio8	$0	$320,000	$0	$7,272,000
Richardson1	$0	$0	$0	($11,412,000)
Richardson2	$0	$0	$0	($11,412,000)
Richardson3	$0	$0	$0	($11,412,000)
Richardson5	$0	$0	$0	($8,076,000)
Richardson6	$0	$0	$0	$88,651,544
Riley1	$0	$0	$0	$1,169,146,954
Riley2	$0	$0	$0	$1,169,146,954
Riley3	$0	$0	$0	($71,300,000)
Riley4	$0	$0	$0	($14,729,000)
Riley5	$0	$0	$0	($4,711,702)
Riley6	$0	$0	$0	$916,401,040
Schofield1	$0	$5,483,000	$0	($99,999,000)
Schofield2	$0	$15,214,000	$0	$55,925,966

Appendix E: Summary Statistics and Correlation Coefficients for Air Force Data

Summary Statistics

Miles (pers)		Equip(T)		Vehicles	
Mean	1,018	Mean	2,870	Mean	442
Standard Error	46	Standard Error	81	Standard Error	45
Median	1,000	Median	3,000	Median	447
Mode	1,000	Mode	3,000	Mode	0
Std Deviation	369	Std Deviation	649	Std Deviation	359
Variance	1.36E+05	Variance	4.21E+05	Variance	1.29E+05
Minimum	209	Minimum	0	Minimum	0
Maximum	2,114	Maximum	3,750	Maximum	1,963
Sum	65,136	Sum	183,671	Sum	28,319
Count	64	Count	64	Count	64
95% Conf. Level	90	95% Conf. Level	159	95% Conf. Level	88

Enlisted		Civilians		Officers	
Mean	2,539	Mean	1,521	Mean	405
Standard Error	184	Standard Error	358	Standard Error	34
Median	2,578	Median	562	Median	346
Mode	1,206	Mode	556	Mode	302
Std Deviation	1,470	Std Deviation	2,867	Std Deviation	274
Variance	2.16E+06	Variance	8.22E+06	Variance	7.49E+04
Minimum	64	Minimum	289	Minimum	33
Maximum	5,576	Maximum	14,463	Maximum	1,413
Sum	162,517	Sum	97,371	Sum	25,935
Count	64	Count	64	Count	64
95% Conf. Level	360	95% Conf. Level	702	95% Conf. Level	67

Facilities (SF)		Acreage		Area Cost Factor	
Mean	2.81E+06	Mean	4,943	Mean	0.98
Standard Error	393,877	Standard Error	979	Standard Error	0.01
Median	2.25E+06	Median	3,570	Median	1.00
Mode	2.64E+06	Mode	0	Mode	1.00
Std Deviation	3.15E+06	Std Deviation	7,835	Std Deviation	0.11
Variance	9.93E+12	Variance	6.14E+07	Variance	1.21E-02
Minimum	262,100	Minimum	0	Minimum	0.77
Maximum	1.49E+07	Maximum	50,999	Maximum	1.26
Sum	1.80E+08	Sum	316,334	Sum	63
Count	64	Count	64	Count	64
95% Conf. Level	771,985	95% Conf. Level	1,920	95% Conf. Level	0.03

RPMA		Comm Budget		Base Ops Cost	
Mean	1.04E+07	Mean	1,370,844	Mean	4,922,672
Standard Error	953,199	Standard Error	152,300	Standard Error	396,345
Median	8,917,000	Median	1,000,000	Median	5,651,000
Mode	1.36E+07	Mode	700,000	Mode	0
Std Deviation	7,625,595	Std Deviation	1,218,397	Std Deviation	3,170,763
Variance	5.815E+13	Variance	1.48E+12	Variance	1.005E+13
Minimum	25,000	Minimum	11,000	Minimum	0
Maximum	4.77E+07	Maximum	6,100,000	Maximum	1.32E+07
Sum	6.65E+08	Sum	8.77E+07	Sum	3.15E+08
Count	64	Count	64	Count	64
95% Conf. Level	1,868,234	95% Conf. Level	298,501	95% Conf. Level	776,821

MFH Budget		CHAMPUS(in)		CHAMPUS (out)	
Mean	2,812,500	Mean	-1905000	Mean	-141128
Standard Error	333,649	Standard Error	574,525	Standard Error	73,811
Median	2,700,000	Median	0	Median	0
Mode	0	Mode	0	Mode	0
Std Deviation	2,669,195	Std Deviation	4,596,200	Std Deviation	590,488
Variance	7.125E+12	Variance	2.113E+13	Variance	3.49E+11
Minimum	0	Minimum	-2.16E+07	Minimum	-3.85E+06
Maximum	8,300,000	Maximum	2.00E+07	Maximum	2,168,400
Sum	1.80E+08	Sum	-1.22E+08	Sum	-9.03E+06
Count	64	Count	64	Count	64
95% Conf. Level	653,940	95% Conf. Level	1,126,047	95% Conf. Level	144,667

One Time Move		Prop. Trans		Net Constr.	
Mean	2.21E+07	Mean	-716158	Mean	1.53E+08
Standard Error	7,866,962	Standard Error	306,277	Standard Error	2.46E+07
Median	1,065,000	Median	0	Median	1.05E+08
Mode	0	Mode	0	Mode	0
Std Deviation	6.29E+07	Std Deviation	2,450,214	Std Deviation	1.97E+08
Variance	3.961E+15	Variance	6.004E+12	Variance	3.88E+16
Minimum	0	Minimum	-11932000	Minimum	-3.63E+08
Maximum	3.32E+08	Maximum	0	Maximum	7.52E+08
Sum	1.41E+09	Sum	-4.58E+07	Sum	9.77E+09
Count	64	Count	64	Count	64
95% Conf. Level	15,418,939	95% Conf. Level	600,291	95% Conf. Level	4.82E+07

Correlation Coefficients (18 variables + 3 dummies)

KEY

bold (shaded) $0.800 \le |r| \le 0.999$ **bold** (boxed) $0.500 \le |r| \le 0.799$

	DI	*EQ*	*VH*	*OF*	*EN*	*CI*	*FA*	*AC*	*AR*	*RP*	*CO*
DI	1.000										
EQ	-0.204	1.000									
VH	-0.120	0.191	1.000								
OF	-0.173	-0.068	**0.647**	1.000							
EN	-0.189	0.019	**0.633**	**0.888**	1.000						
CI	0.040	-0.231	-0.002	0.367	0.296	1.000					
FA	-0.005	-0.161	0.208	0.574	**0.546**	**0.937**	1.000				
AC	-0.081	0.002	0.164	0.373	0.389	0.062	0.216	1.000			
AR	-0.105	0.146	-0.138	-0.203	-0.206	-0.156	-0.106	-0.109	1.000		
RP	-0.114	-0.062	0.203	**0.548**	**0.588**	**0.807**	**0.895**	0.245	-0.079	1.000	
CO	0.035	-0.072	0.193	0.320	0.239	**0.521**	**0.513**	0.093	-0.203	0.417	1.000
BO	-0.235	0.048	0.421	**0.609**	**0.663**	0.329	**0.544**	0.383	-0.019	**0.565**	0.390
MF	-0.050	0.004	**0.544**	**0.641**	**0.732**	0.108	0.388	0.422	0.079	0.464	0.084
IP	0.361	-0.366	-0.289	-0.351	-0.451	0.287	0.096	-0.192	-0.211	-0.099	0.178
OP	0.195	-0.227	-0.327	-0.156	-0.226	0.277	0.138	-0.017	-0.169	-0.040	0.225
OT	0.158	-0.275	-0.051	0.310	0.196	**0.885**	**0.825**	0.002	-0.098	**0.665**	0.433
PT	0.194	-0.066	-0.130	-0.404	-0.336	0.008	-0.074	-0.408	0.152	-0.123	0.072
NC	-0.193	-0.154	0.365	**0.699**	**0.718**	**0.510**	**0.666**	0.339	-0.041	**0.643**	0.153
RV	-0.021	0.087	-0.322	-0.358	-0.535	-0.148	-0.286	-0.224	0.070	-0.269	0.159
NG	-0.026	0.107	-0.392	-0.472	-0.468	-0.202	-0.374	-0.266	0.086	-0.415	-0.320
DP	0.212	-0.435	-0.104	0.201	0.116	**0.889**	**0.797**	-0.017	-0.129	**0.616**	0.469

	BO	*MF*	*IP*	*OP*	*OT*	*PT*	*NC*	*RV*	*NG*	*DP*
DI										
EQ										
VH										
OF										
EN										
CI										
FA										
AC										
AR										
RP										
CO										
BO	1.000									
MF	**0.638**	1.000								
IP	-0.269	-0.315	1.000							
OP	-0.118	-0.279	0.422	1.000						
OT	0.255	0.075	0.294	0.214	1.000					
PT	-0.228	-0.222	0.436	0.028	0.081	1.000				
NC	**0.563**	0.487	-0.298	0.005	0.433	-0.369	1.000			
RV	-0.107	-0.457	0.180	0.104	-0.151	0.127	-0.336	1.000		
NG	**-0.728**	**-0.562**	0.221	0.127	-0.182	0.156	-0.415	-0.228	1.000	
DP	0.189	0.007	0.369	0.276	**0.928**	0.095	0.348	-0.138	-0.170	1.000

Correlation Coefficients (16 variables + 3 dummies)

[*TP* = *OF* + *EN* + *CI*]

KEY

bold $0.800 \le |r| \le 0.999$ **bold** $0.500 \le |r| \le 0.799$

	DI	*EQ*	*VH*	*TP*	*FA*	*AC*	*AR*	*RP*	*CO*	*BO*
DI	1.000									
EQ	-0.204	1.000								
VH	-0.120	0.191	1.000							
TP	-0.056	-0.173	0.292	1.000						
FA	-0.005	-0.161	0.208	**0.966**	1.000					
AC	-0.081	0.002	0.164	0.225	0.216	1.000				
AR	-0.105	0.146	-0.138	-0.214	-0.106	-0.109	1.000			
RP	-0.114	-0.062	0.203	**0.882**	**0.895**	0.245	-0.079	1.000		
CO	0.035	-0.072	0.193	**0.512**	**0.513**	0.093	-0.203	0.417	1.000	
BO	-0.235	0.048	0.421	**0.553**	**0.544**	0.383	-0.019	**0.565**	0.390	1.000
MF	-0.050	0.004	**0.544**	0.414	0.388	0.422	0.079	0.464	0.084	**0.638**
IP	0.361	-0.366	-0.289	0.016	0.096	-0.192	-0.211	-0.099	0.178	-0.269
OP	0.195	-0.227	-0.327	0.111	0.138	-0.017	-0.169	-0.040	0.225	-0.118
OT	0.158	-0.275	-0.051	**0.771**	**0.825**	0.002	-0.098	**0.665**	0.433	0.255
PT	0.194	-0.066	-0.130	-0.154	-0.074	-0.408	0.152	-0.123	0.072	-0.228
NC	-0.193	-0.154	0.365	**0.718**	**0.666**	0.339	-0.041	**0.643**	0.153	**0.563**
RV	-0.026	0.107	-0.392	-0.370	-0.374	-0.266	0.086	-0.415	-0.320	**-0.728**
NG	-0.021	0.087	-0.322	-0.347	-0.286	-0.224	0.070	-0.269	0.159	-0.107
DP	0.212	-0.435	-0.104	**0.735**	**0.797**	-0.017	-0.129	**0.616**	0.469	0.189

	MF	*IP*	*OP*	*OT*	*PT*	*NC*	*RV*	*NG*	*DP*
DI									
EQ									
VH									
TP									
FA									
AC									
AR									
RP									
CO									
BO									
MF	1.000								
IP	-0.315	1.000							
OP	-0.279	0.422	1.000						
OT	0.075	0.294	0.214	1.000					
PT	-0.222	0.436	0.028	0.081	1.000				
NC	0.487	-0.298	0.005	0.433	-0.369	1.000			
RV	**-0.562**	0.221	0.127	-0.182	0.156	-0.415	1.000		
NG	-0.457	0.180	0.104	-0.151	0.127	-0.336	-0.228	1.000	
DP	0.007	0.369	0.276	**0.928**	0.095	0.348	-0.170	-0.138	1.000

Correlation Coefficients (12 variables + 3 dummies)

[Eliminate FA, AR, and OT.]
[CH = IP + OP]

KEY

bold $0.800 \leq |r| \leq 0.999$ **bold** $0.500 \leq |r| \leq 0.799$

	DI	*EQ*	*VH*	*TP*	*AC*	*RP*	*CO*	*BO*
DI	1.000							
EQ	-0.204	1.000						
VH	-0.120	0.191	1.000					
TP	-0.056	-0.173	0.292	1.000				
AC	-0.081	0.002	0.164	0.225	1.000			
RP	-0.114	-0.062	0.203	**0.882**	0.245	1.000		
CO	0.035	-0.072	0.193	**0.512**	0.093	0.417	1.000	
BO	-0.235	0.048	0.421	**0.553**	0.383	**0.565**	0.390	1.000
MF	-0.050	0.004	**0.544**	0.414	0.422	0.464	0.084	**0.638**
CH	0.364	-0.373	-0.312	0.029	-0.183	-0.098	0.195	-0.268
PT	0.194	-0.066	-0.130	-0.154	-0.408	-0.123	0.072	-0.228
NC	-0.193	-0.154	0.365	**0.718**	0.339	**0.643**	0.153	**0.563**
RV	-0.026	0.107	-0.392	-0.370	-0.266	-0.415	-0.320	-0.728
NG	-0.021	0.087	-0.322	-0.347	-0.224	-0.269	0.159	-0.107
DP	0.212	-0.435	-0.104	**0.735**	-0.017	**0.616**	0.469	0.189

	MF	*CH*	*PT*	*NC*	*RV*	*NG*	*DP*
DI							
EQ							
VH							
TP							
AC							
RP							
CO							
BO							
MF	1.000						
CH	-0.331	1.000					
PT	-0.222	0.415	1.000				
NC	0.487	-0.280	-0.369	1.000			
RV	**-0.562**	0.224	0.156	-0.415	1.000		
NG	-0.457	0.182	0.127	-0.336	-0.228	1.000	
DP	0.007	0.382	0.095	0.348	-0.170	-0.138	1.000

Appendix F: SAS Instruction Files

Reading Data

The following instruction set reads the data in from an ASCII text file into an internal SAS database.

```
OPTION LINESIZE = 76;
TITLE ' Analysis of the Cost Of Base Realignment Action (COBRA)';
TITLE1 'cost output against selected inputs';
DATA COBRADAT;
  INFILE 'BASEDATA.DAT';
  INPUT BASE$ NPV DI EQ VH OF EN CI FA AC AR RP CO BO MF IP OP OT PT
NC RV NG DP MY;
  TOTPER = EN + CI + OF;
PROC PRINT DATA = COBRADAT;
   TITLE 'FINAL DATA FROM INDIVIDUAL BASES';
   ID BASE;
RUN;
```

Correlational Analysis of Air Force Data Set

This instruction set completes a correlational analysis of the Air Force data set -- active duty, AFRes, ANG, Depot.

```
OPTION LINESIZE = 76;
TITLE 'the Correlation Matrix derived from the AF data file';
DATA COBRADAT;
  INFILE 'AFDATA.DAT';
  INPUT BASE$ NPV DI EQ VH OF EN CI FA AC AR RP CO BO MF IP OP OT PT
NC RV NG DP MY;
PROC CORR DATA = COBRADAT RANK;
  VAR DI EQ VH OF EN CI FA AC AR RP CO BO MF IP OP OT PT NC RV NG DP MY;
RUN;
```

Model Generation

This instruction set reads in the data and employs the stepwise selection and maximum R^2 procedures. Stepwise produces a single "best" model according to its own methodologies. Maximum R^2 generates the "best" model for each number of variables (i.e. the best 1 variable model, the best 2 variable, ..., the best *m* variable model with *m* being the total number of variables to be regressed). Each selection procedure was executed with and without an intercept term.

```
TITLE ' MODEL SELECTION OUTPUT ';

DATA COBRADAT;

  INFILE 'AFDATA.DAT';

  INPUT BASE$ NPV DI EQ VH OF EN CI FA AC AR RP CO BO MF
IP OP OT PT NC RV NG DP MY;

  TOTPER = EN + OF + CI;

PROC REG DATA = COBRADAT;

  MODEL NPV = DI EQ VH TOTPER AC AR RP CO BO MF IP OP PT
NC RV NG DP;

  MODEL NPV = DI EQ VH TOTPER AC AR RP CO BO MF IP OP PT
NC RV NG DP/NOINT;

  MODEL NPV = DI EQ VH TOTPER AC RP CO BO MF IP OP PT NC
RV NG DP /SELECTION = STEPWISE SLE = .30 SLS = .10;

  MODEL NPV   DI EQ VH TOTPER AC RP CO BO MF IP OP PT NC
RV NG DP /SELECTION = STEPWISE SLE = .3 SLS = .05;

  MODEL NPV = DI EQ VH TOTPER AC RP CO BO MF IP OP PT NC
RV NG DP /SELECTION = MAXR STOP = 13;

  MODEL NPV = DI EQ VH TOTPER AC AR RP CO BO MF IP OP PT
NC RV NG DP /SELECTION = STEPWISE NOINT SLE = .3 SLS = .1;

  MODEL NPV = DI EQ VH TOTPER AC AR RP CO BO MF IP OP PT
NC RV NG DP /SELECTION = STEPWISE NOINT SLE = .3 SLS = .05;

  MODEL NPV = DI EQ VH TOTPER AC AR RP CO BO MF IP OP PT
NC RV NG DP /SELECTION = MAXR NOINT STOP = 13;

RUN;
```

Appendix G: Residual Plots for Model 1

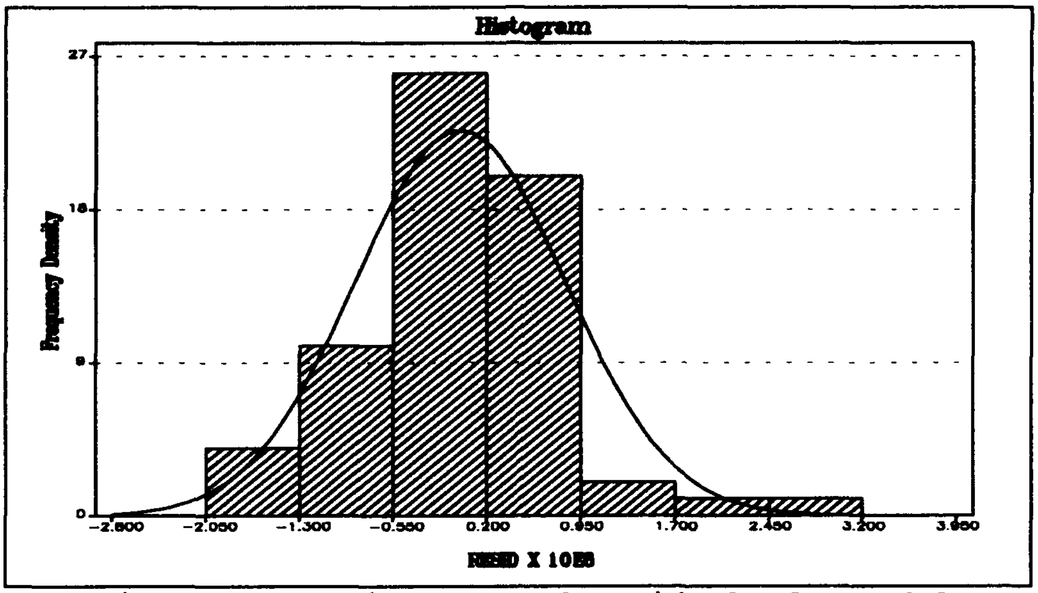


Figure G.1. Histogıam of residuals for Model

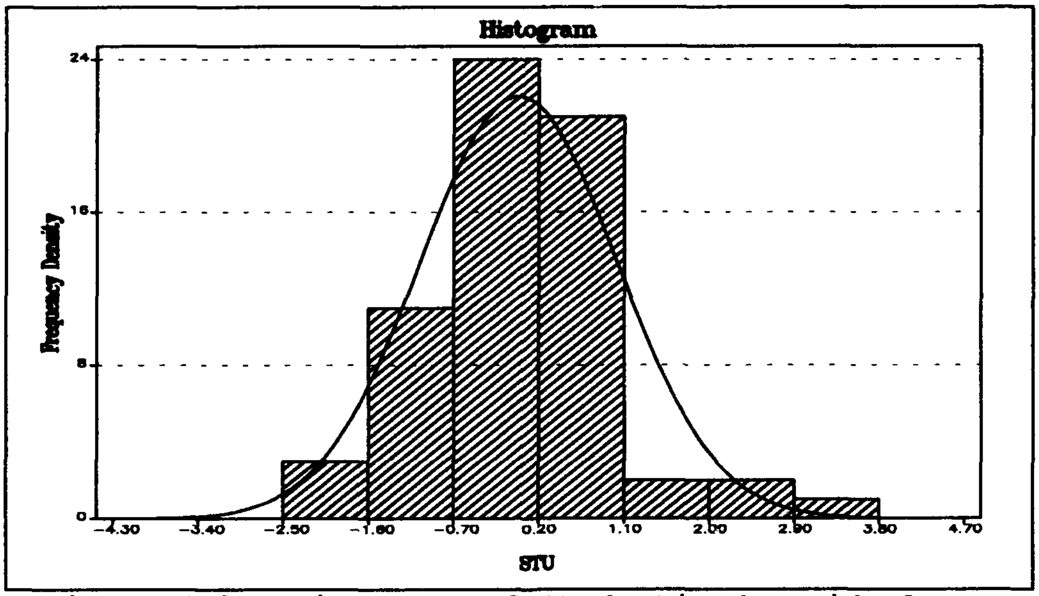


Figure G.2. Histogram of Studentized residuals for Model 1.

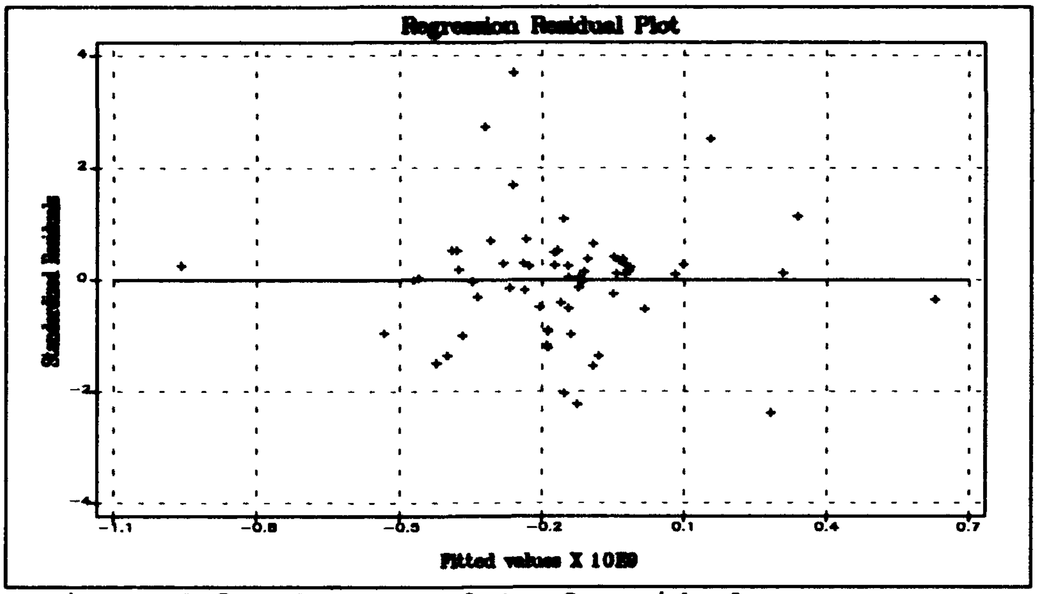


Figure G.3. Scatter Plot of residuals versus predicted values of *NPV* for Model 1.

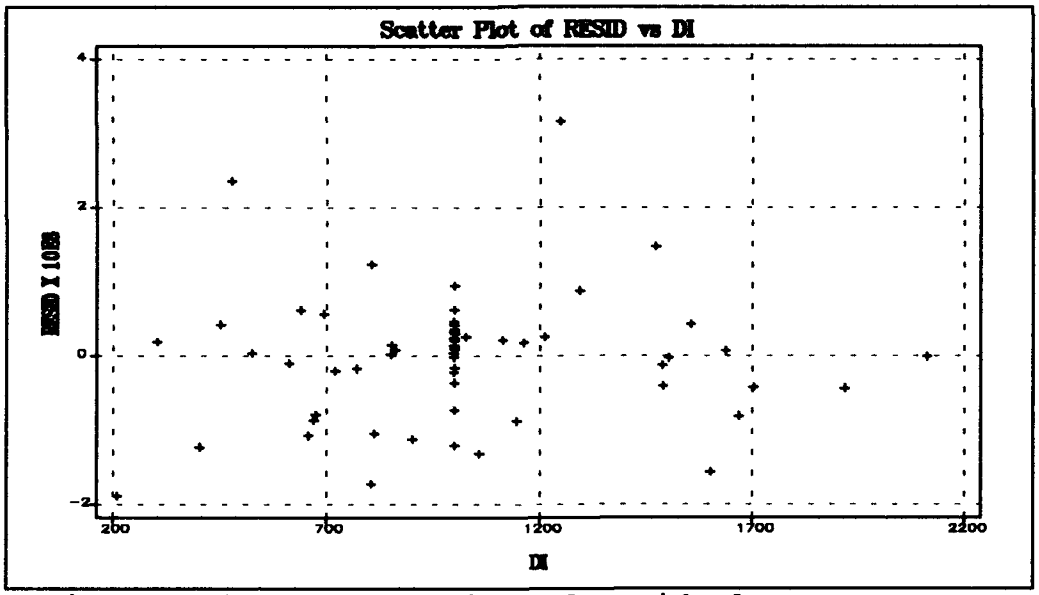


Figure G.4. Scatter plot of residuals versus independent variable *DI* for Model 1.

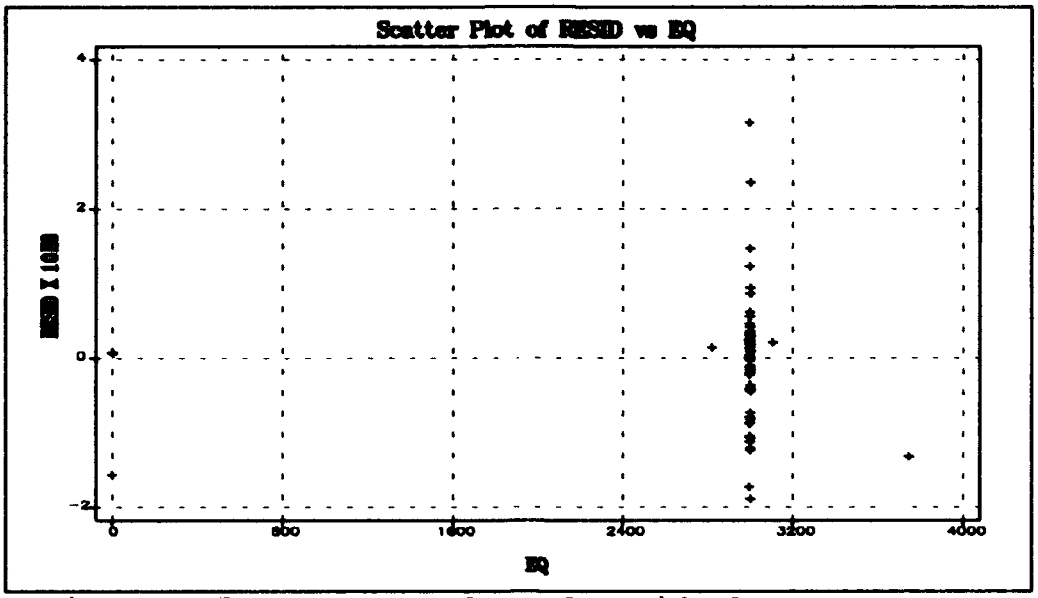


Figure G.5. Scatter plot of residuals versus independent variable *EQ* for Model 1.

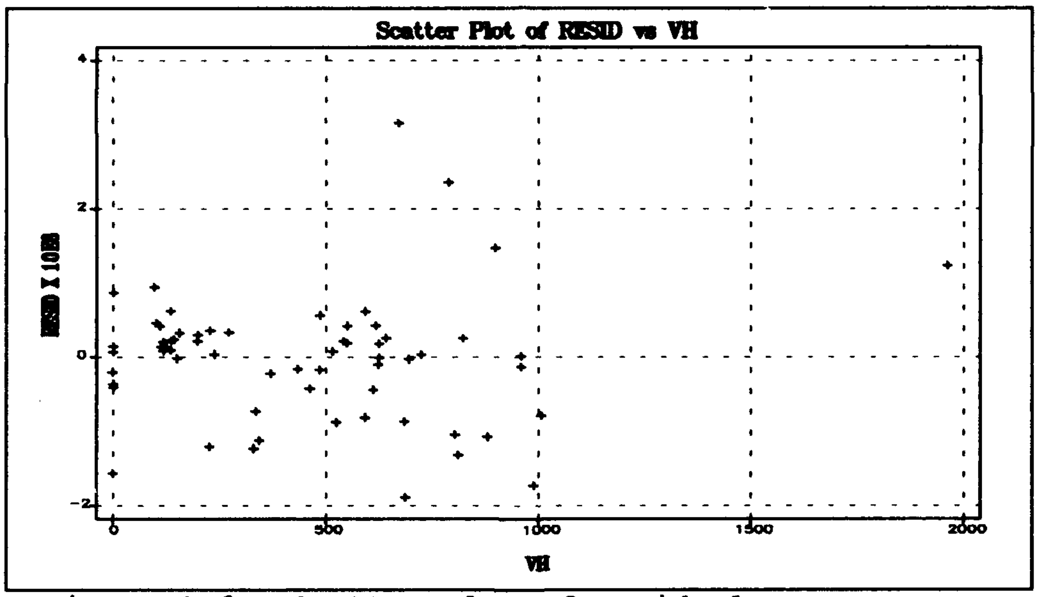


Figure G.6. Scatter plot of residuals versus independent variable *VH* for Model 1.

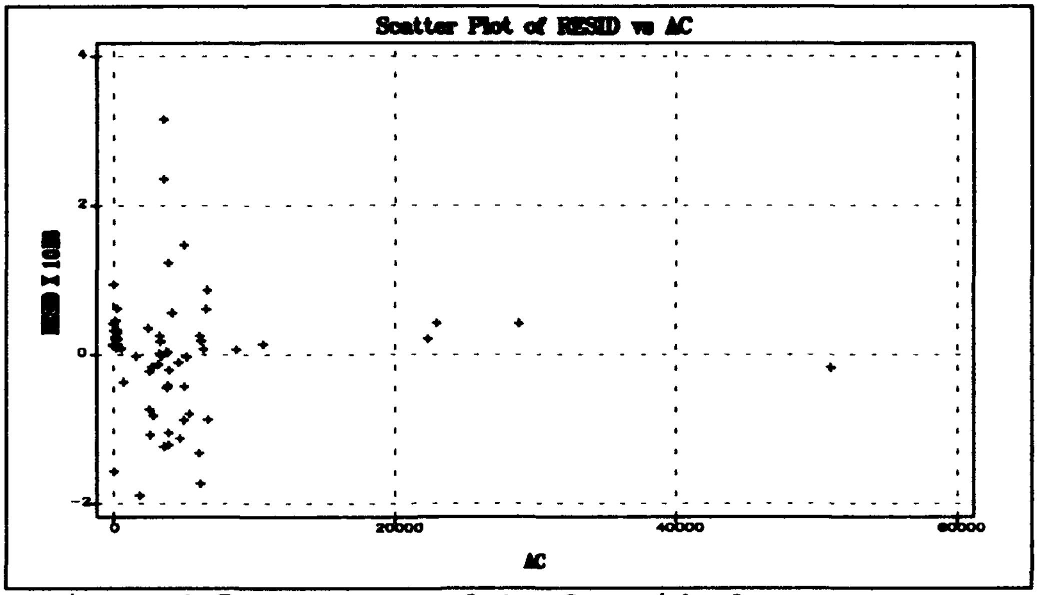


Figure G.7. Scatter plot of residuals versus independent variable *AC* for Model 1.

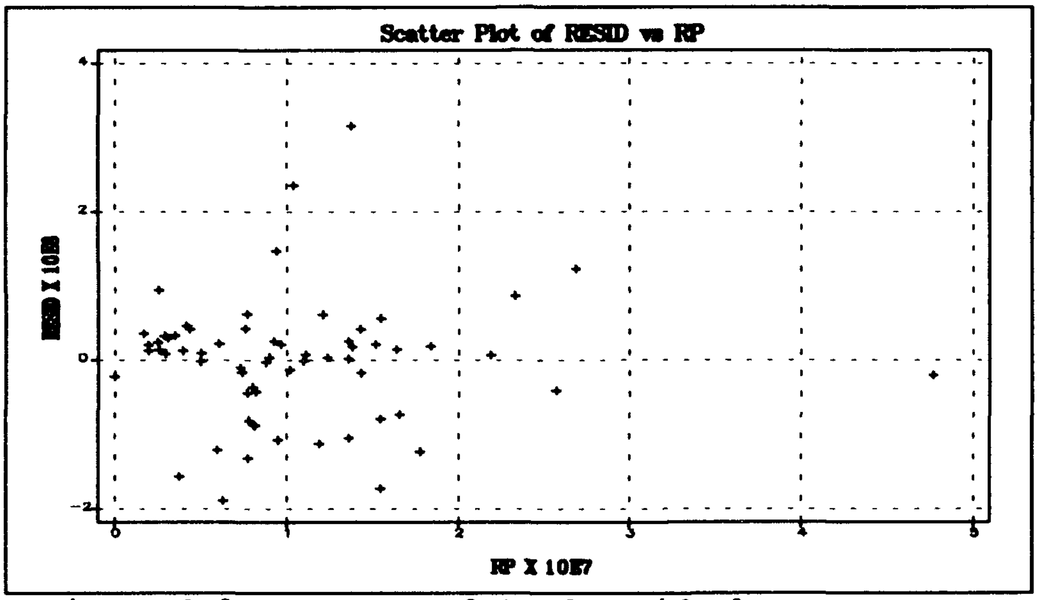


Figure G.8. Scatter plot of residuals versus independent variable *RP* for Model 1.

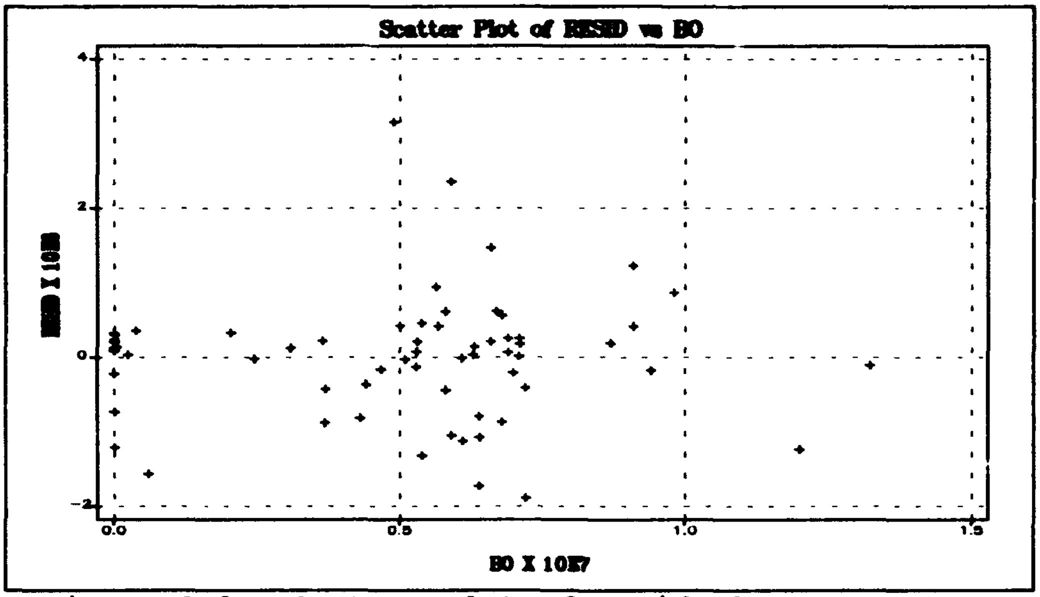


Figure G.9. Scatter plot of residuals versus independent variable *BO* for Model 1.

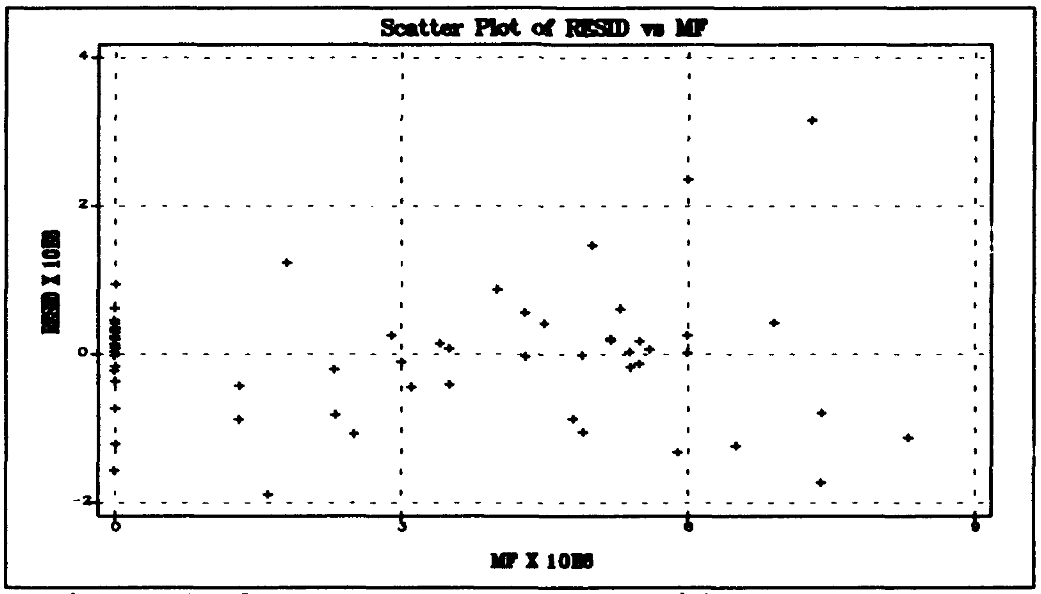
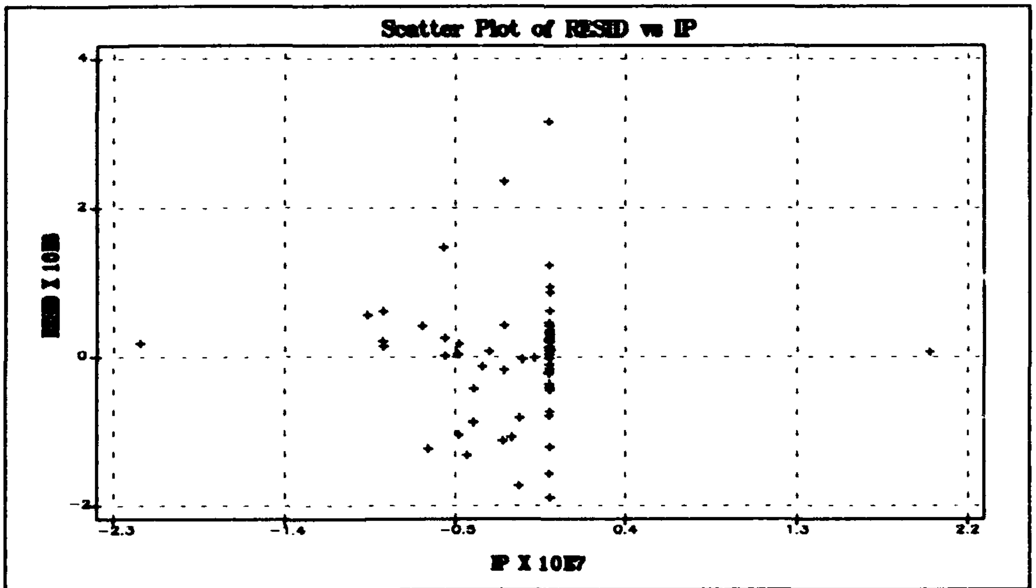


Figure G.10. Scatter plot of residuals versus independent variable *MF* for Model 1.



Figur G.11. Scatter plot of residuals versus independent variable *IP* for Model 1.

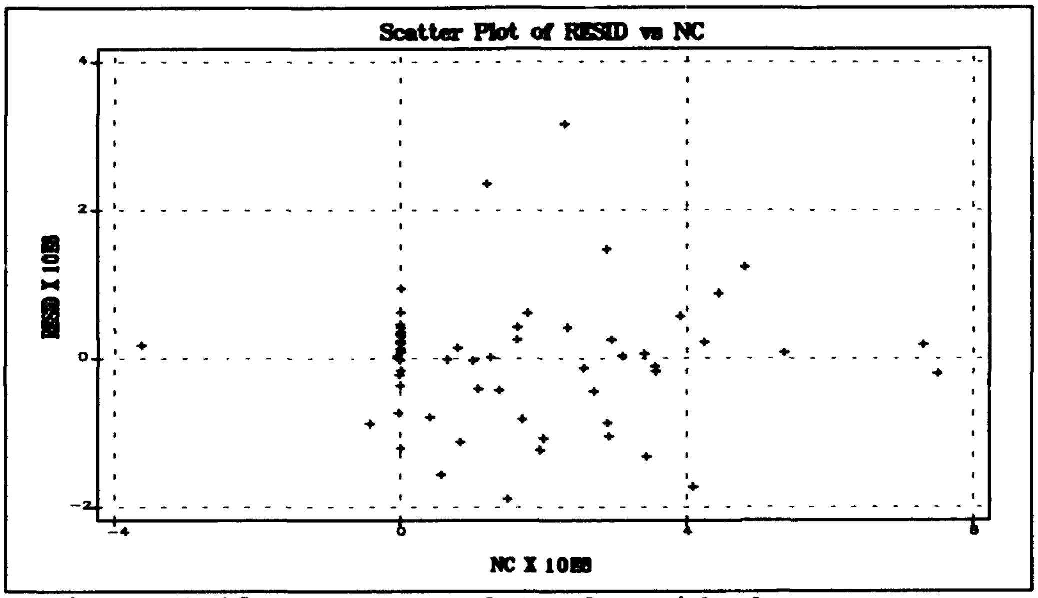


Figure G.12. Scatter plot of residuals versus independent variable *NC* for Model 1.

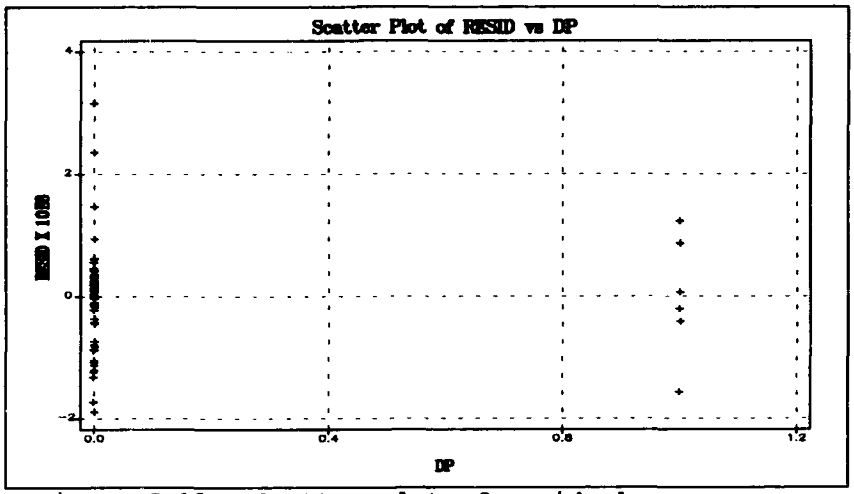


Figure G.13. Scatter plot of residuals versus independent (dummy) variable *DP* for Model 1.

Appendix H: Residual Plots for Model 2

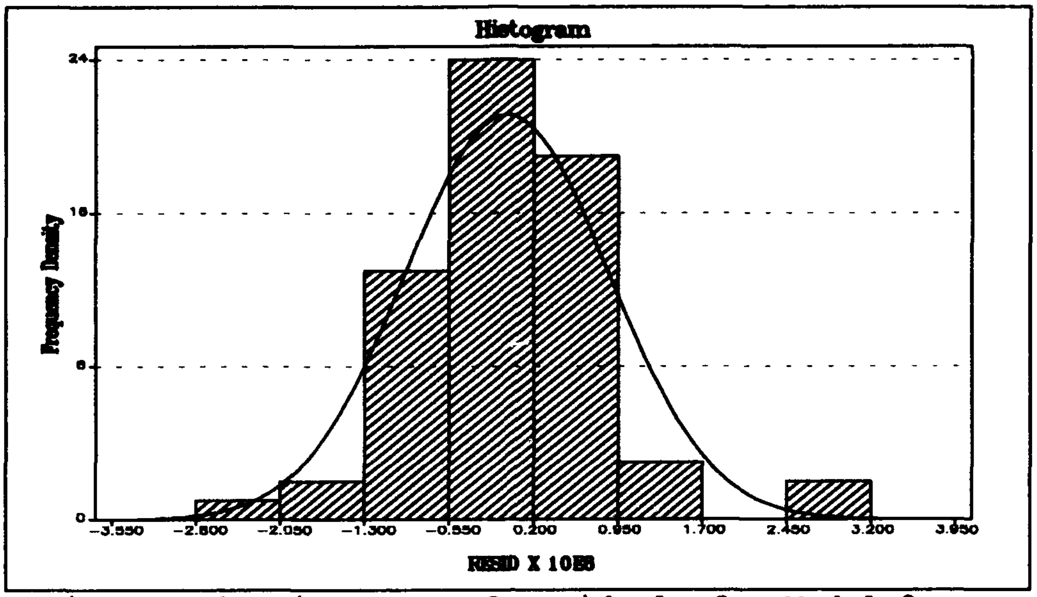


Figure H.1 Histogram of residuals for Model 2.

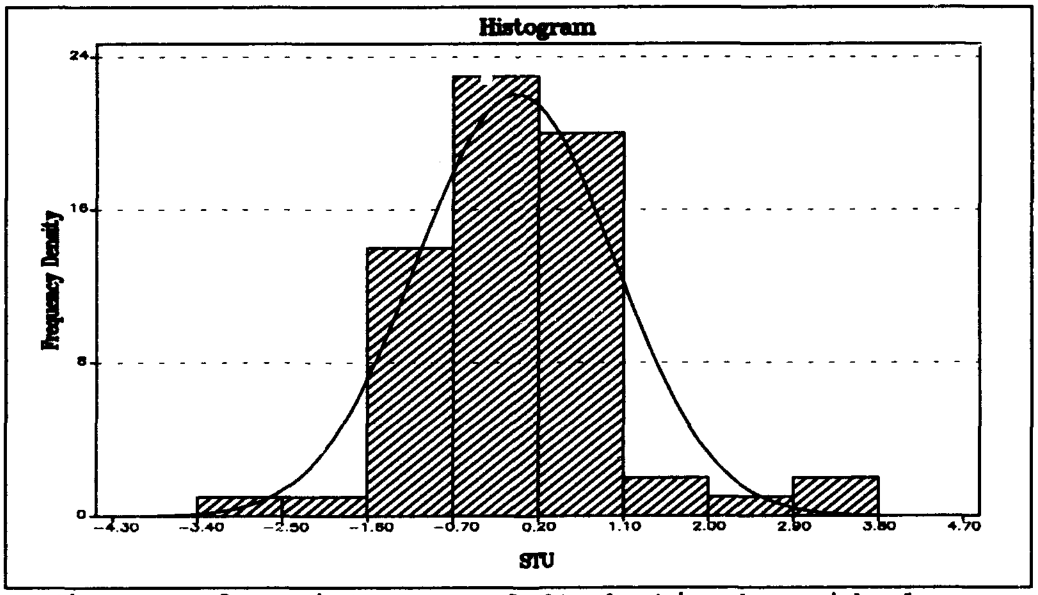


Figure H.2. Histogram of Studentized residuals for Model 2.

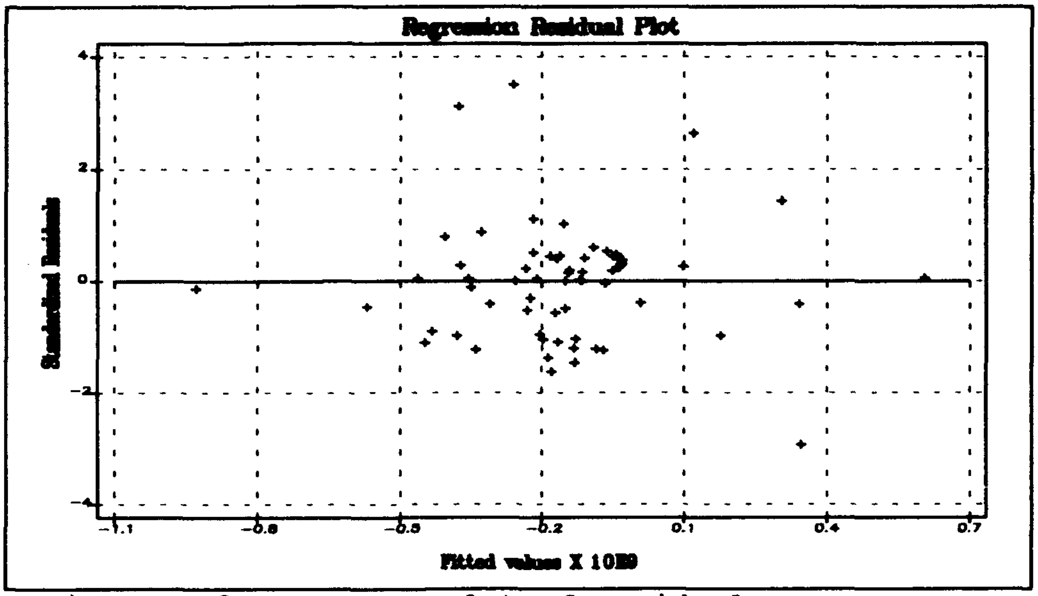


Figure H.3. Scatter plot of residuals versus predicted values of *NPV* for Model 2.

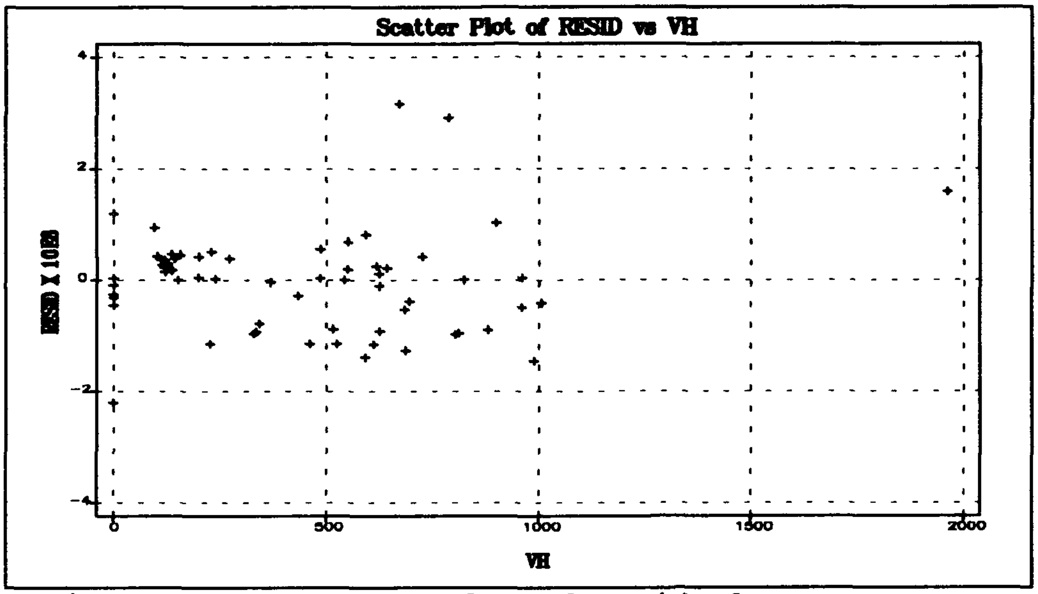


Figure H.4. Scatter plot of residuals versus independent variable *VH* for Model 2.

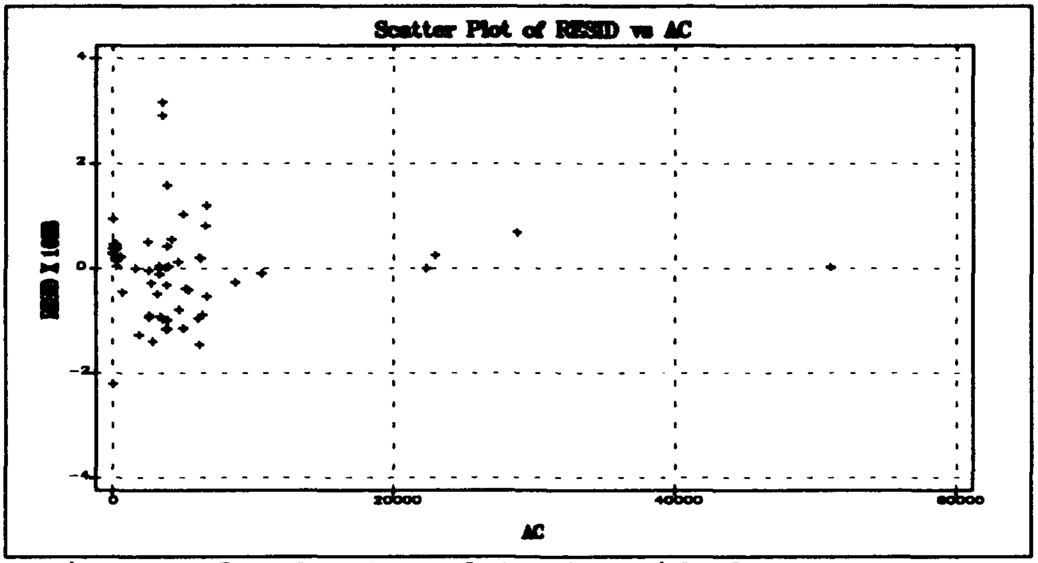


Figure H.5. Scatter plot of residuals versus independent variable *AC* for Model 2.

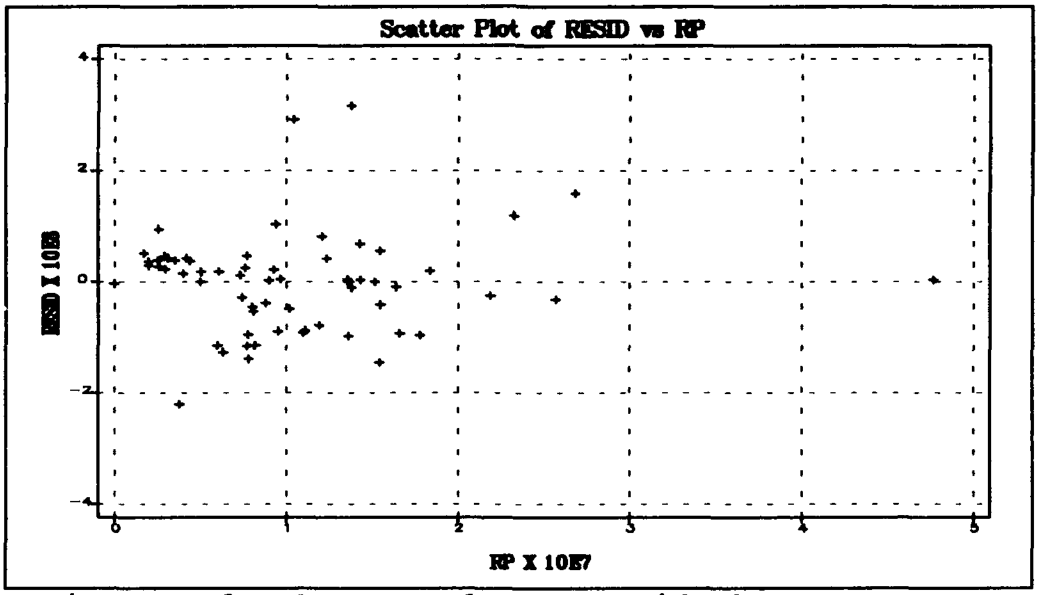


Figure H.6. Scatter plot of residuals versus independent variable *RP* for Model 2.

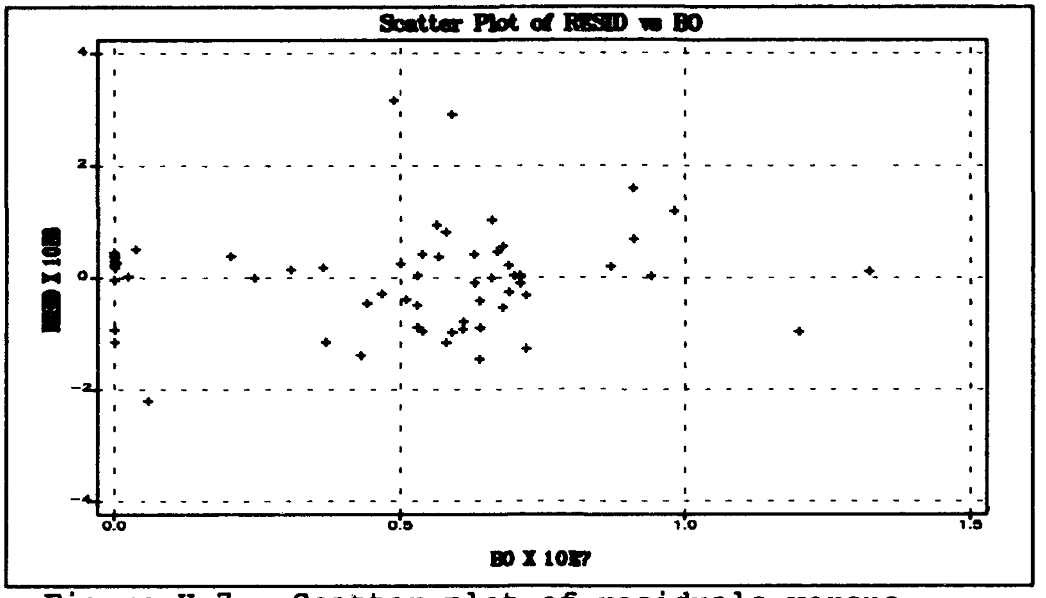


Figure H.7. Scatter plot of residuals versus independent variable *BO* for Model 2.

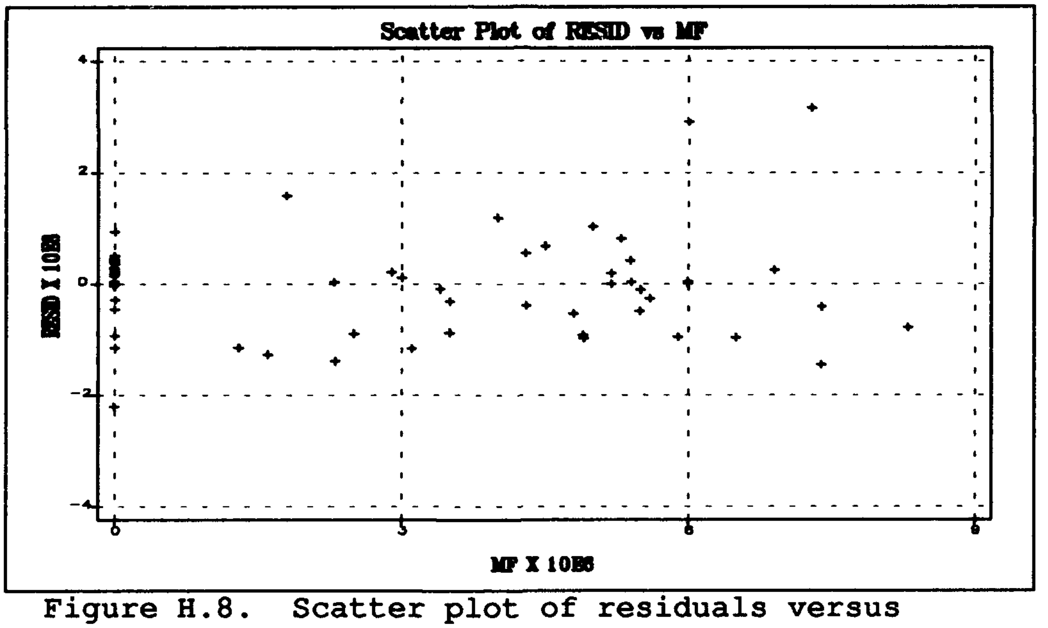


Figure H.8. Scatter plot of residuals versus independent variable *MF* for Model 2.

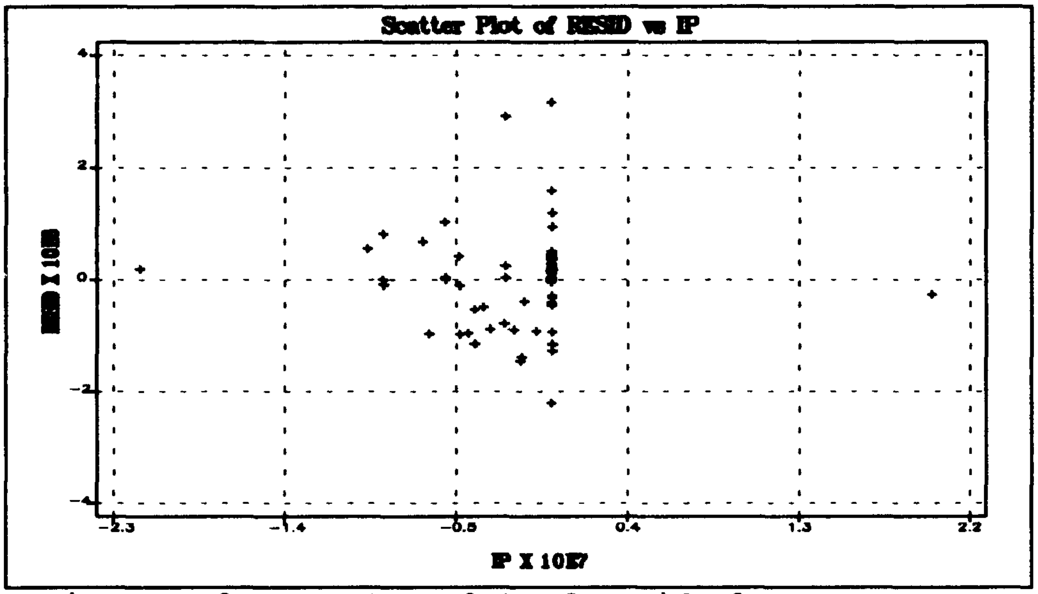


Figure H.9. Scatter plot of residuals versus independent variable *IP* for Model 2.

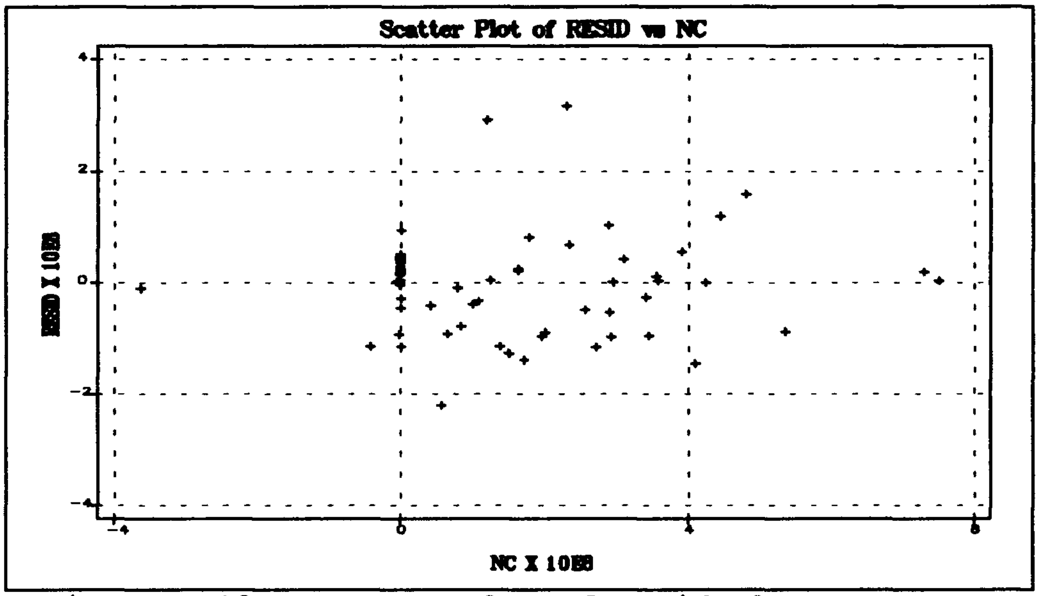


Figure H.10. Scatter plot of residuals versus independent variable *NC* for Model 2.

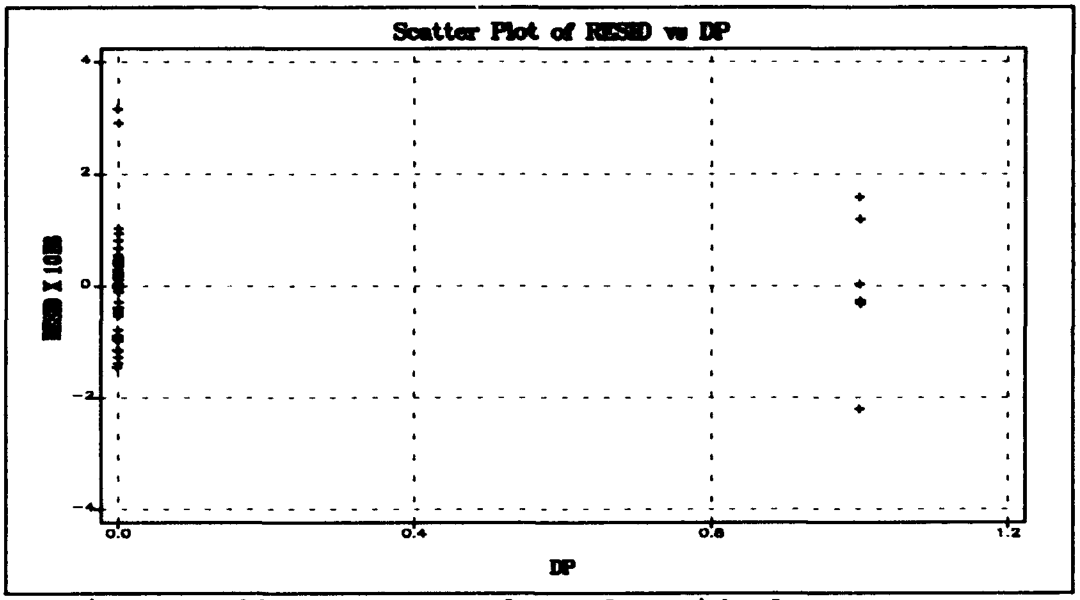


Figure H.11. Scatter plot of residuals versus independent (dummy) variable DP for Model 2.

Appendix I: Residual Plots for Model 3

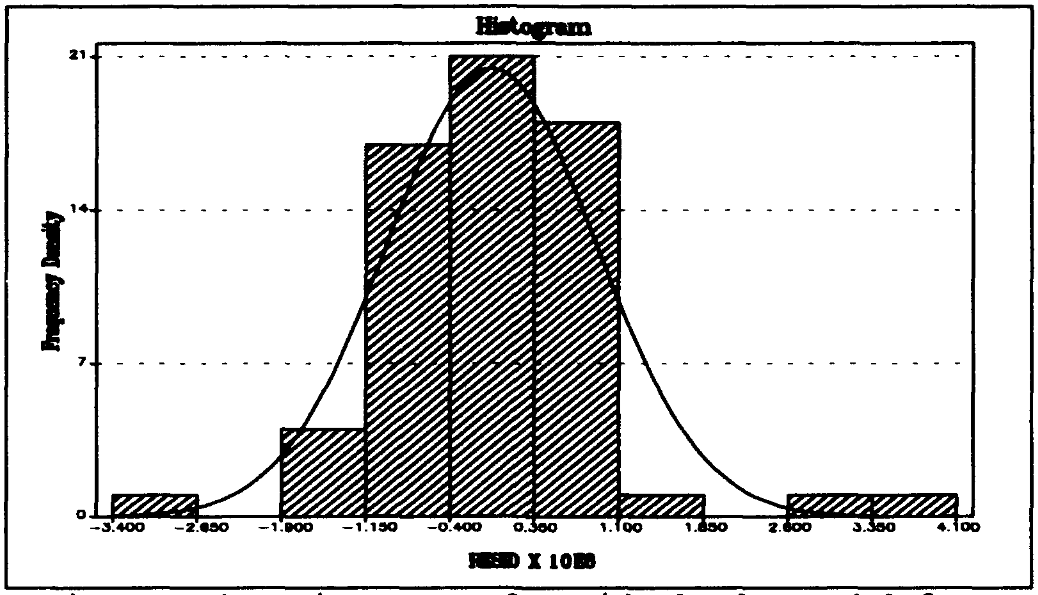


Figure I.1. Histogram of residuals for Model 3.

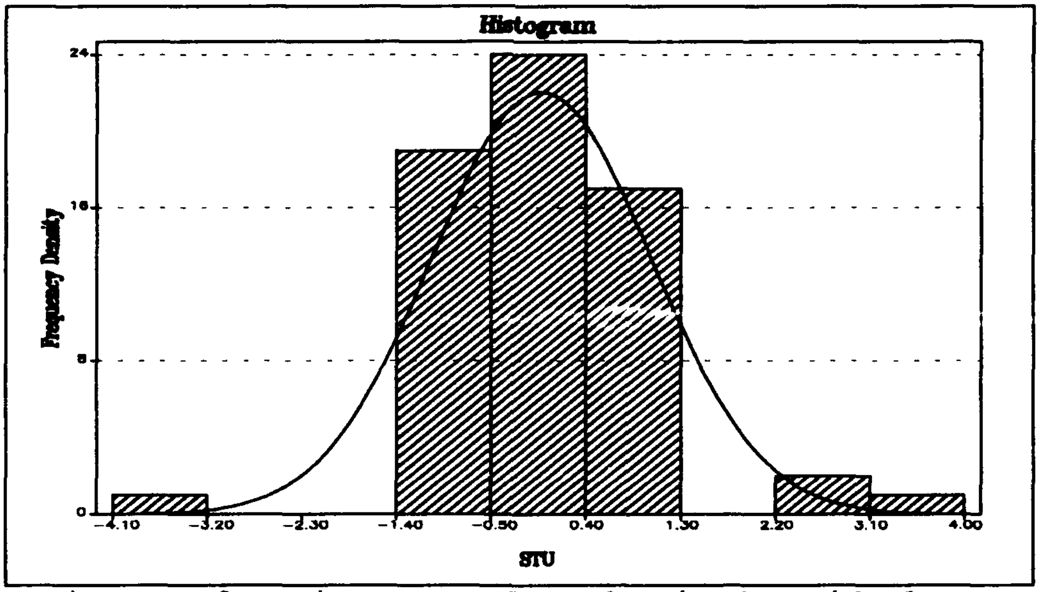


Figure I.2. Histogram of Studentized residuals for Model 3.

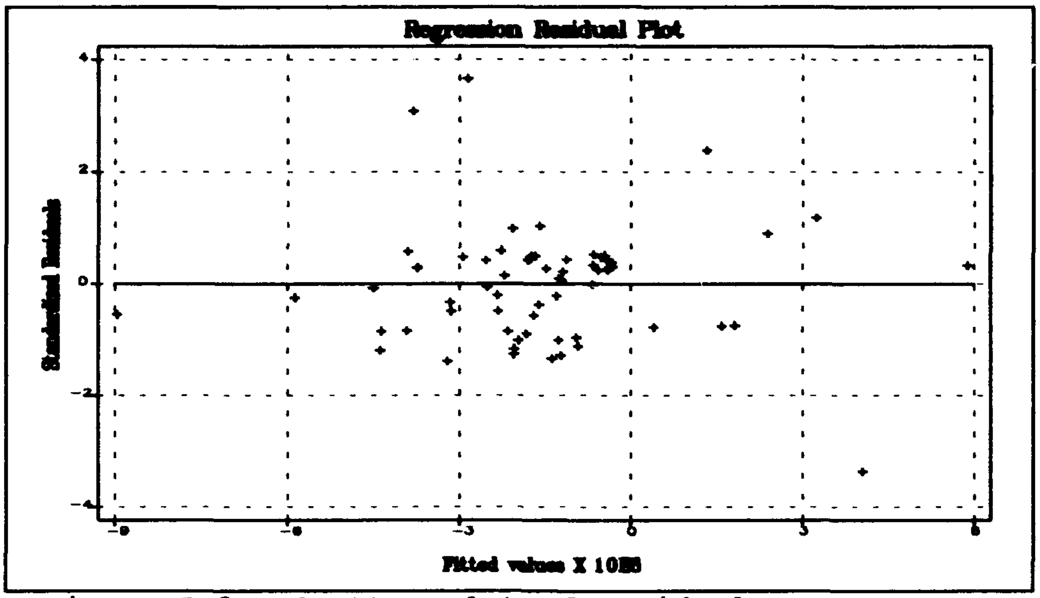


Figure I.3. Scatter plot of residuals versus predicted values of *NPV* for Model 3.

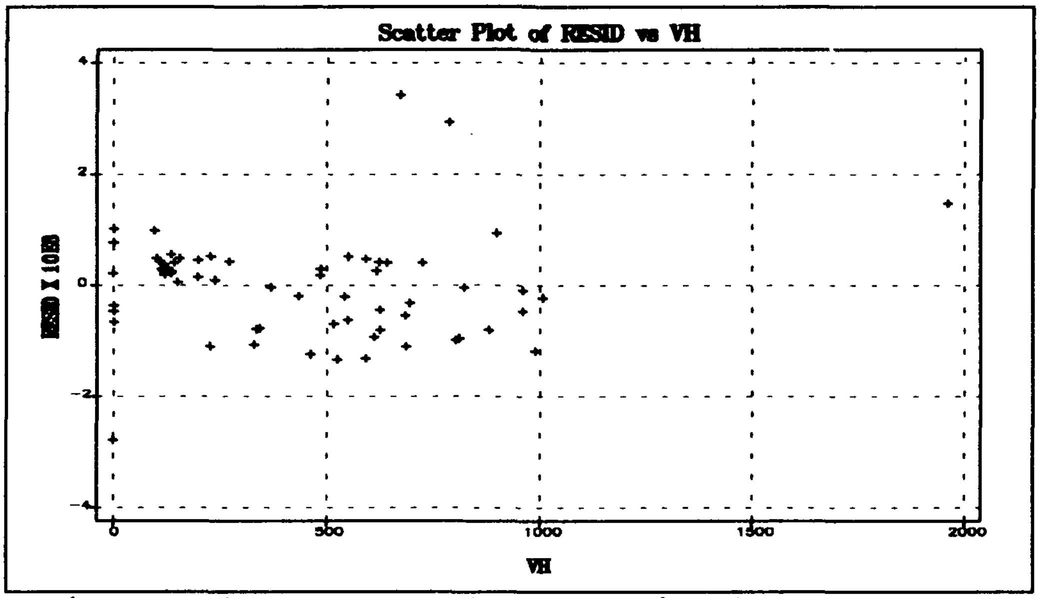


Figure I.4. Scatter plot of residuals versus independent variable VH for Model 3.

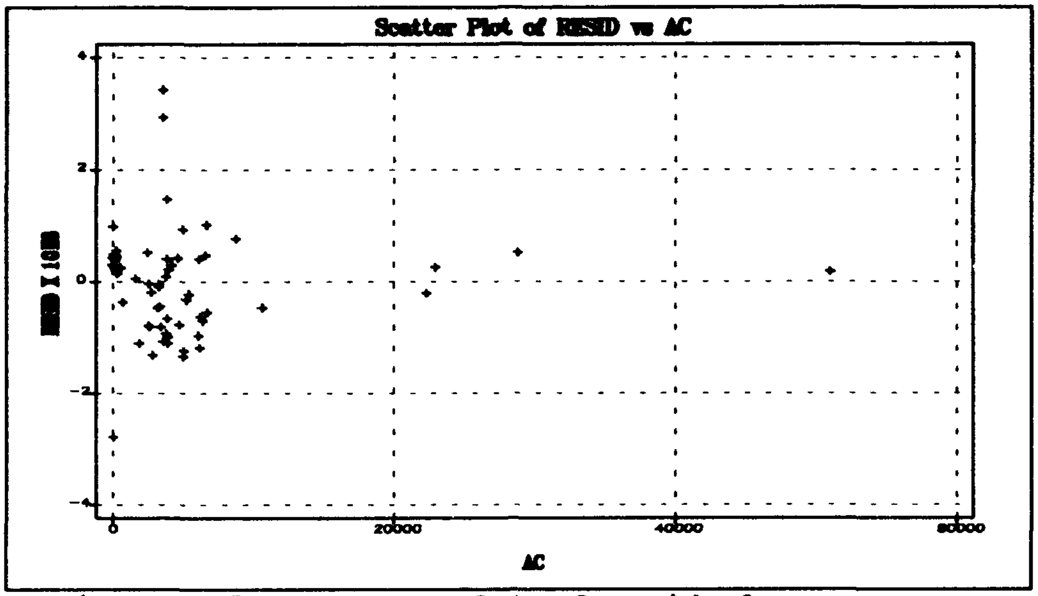


Figure I.5. Scatter plot of residuals versus independent variable *AC* for Model 3.

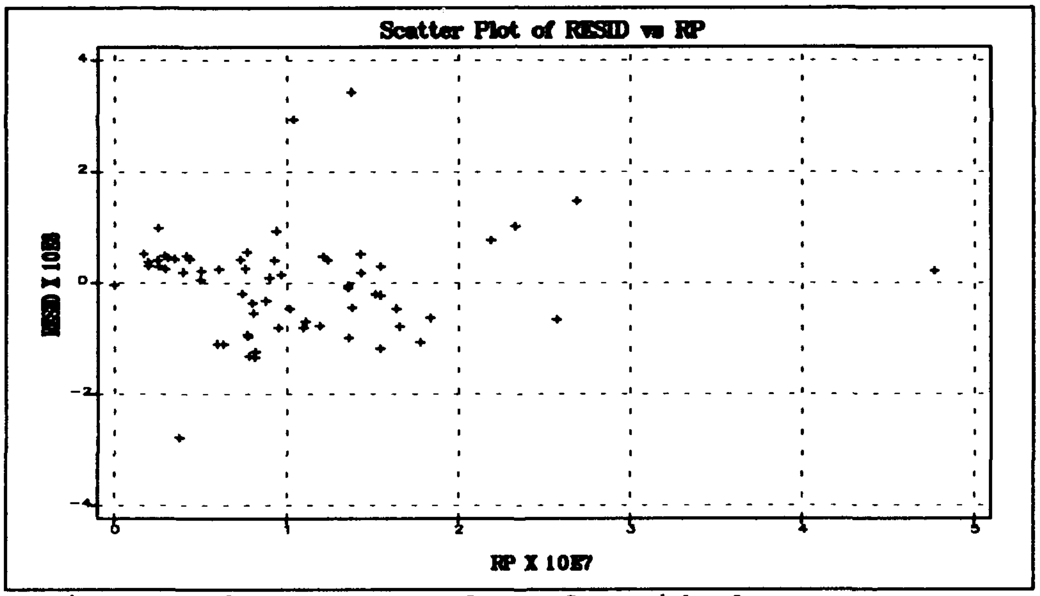


Figure I.6. Scatter plot of residuals versus independent variable *RP* for Model 3.

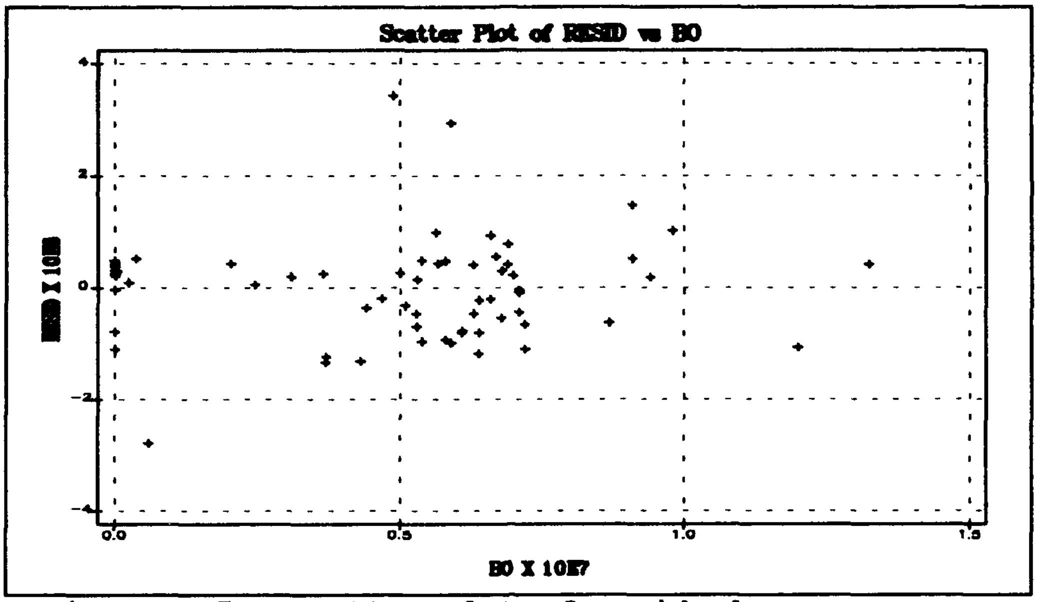


Figure I.7. Scatter plot of residuals versus independent variable *BO* for Model 3.

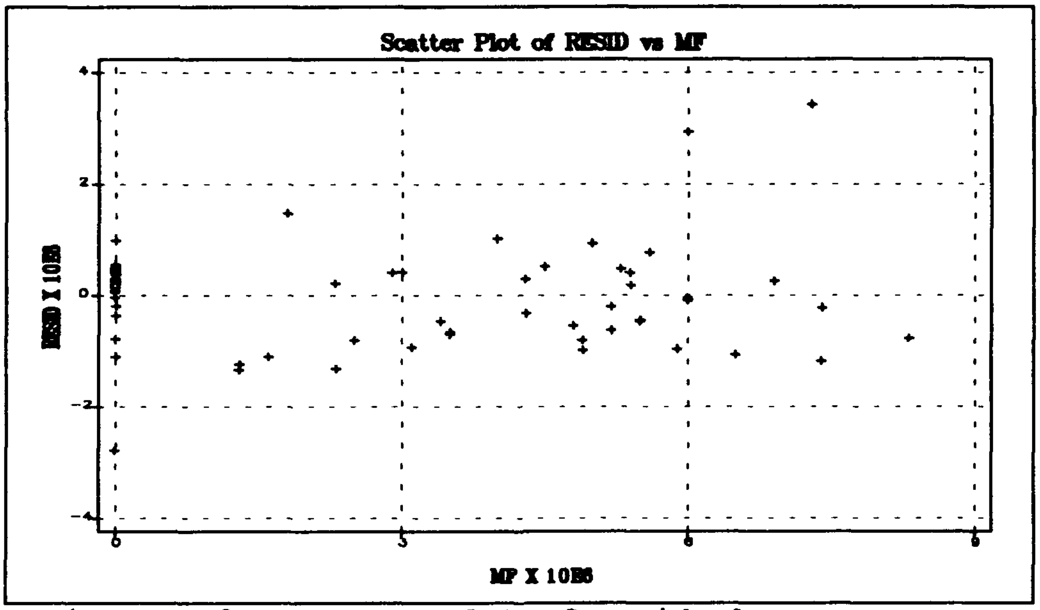


Figure I.8. Scatter plot of residuals versus independent variable *MF* for Model 3.

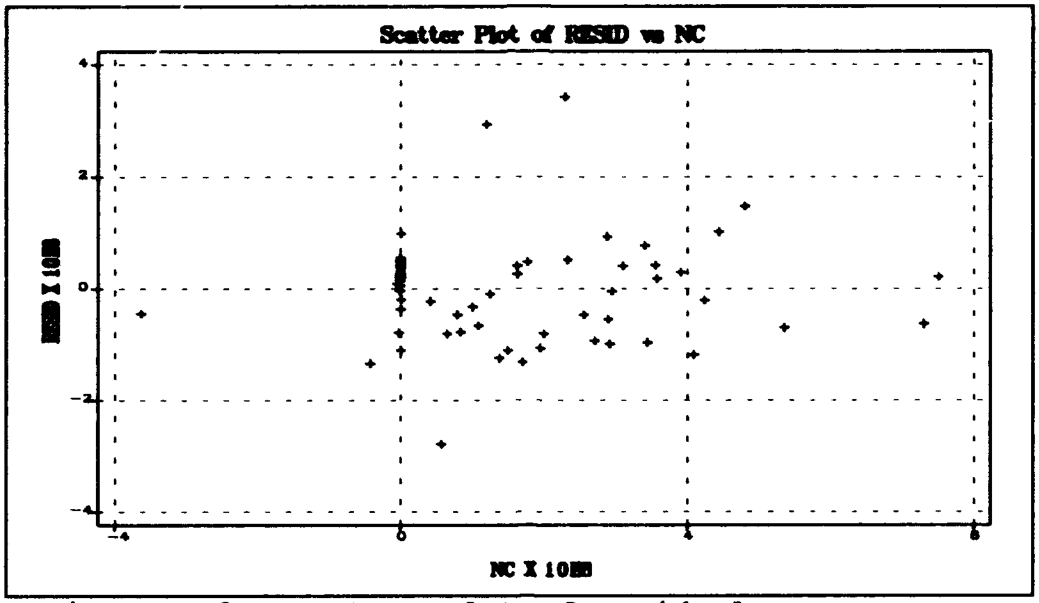


Figure I.9. Scatter plot of residuals versus independent variable *NC* for Model 3.

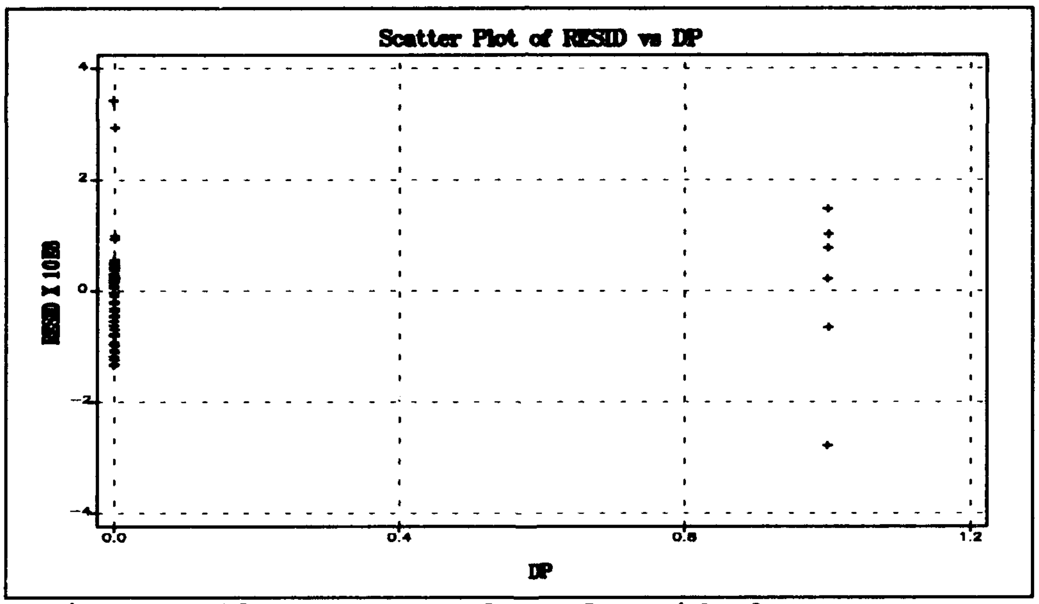


Figure I.10. Scatter plot of residuals versus independent (dummy) variable *DP* for Model 3.

Appendix J: Residual Plots for Model 4

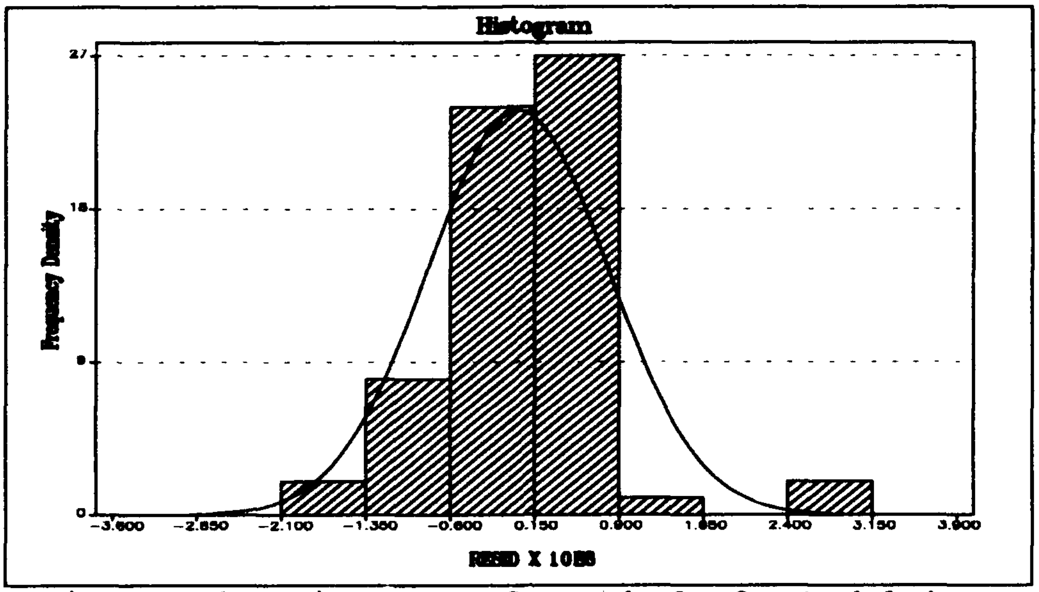


Figure J.1. Histogram of residuals for Model 4.

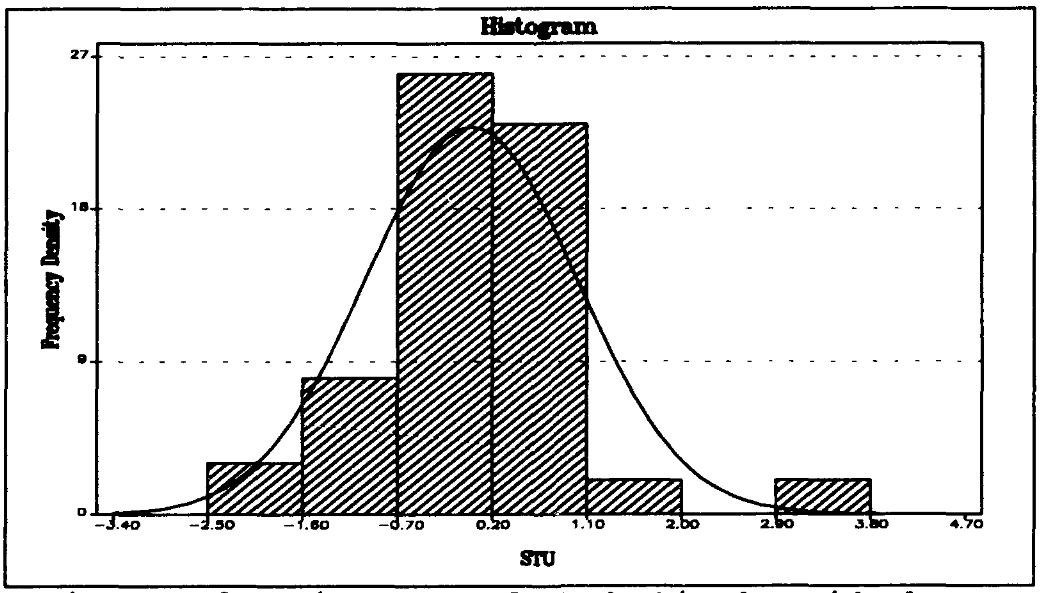


Figure J.2. Histogram of Studentized residuals for Model 4.

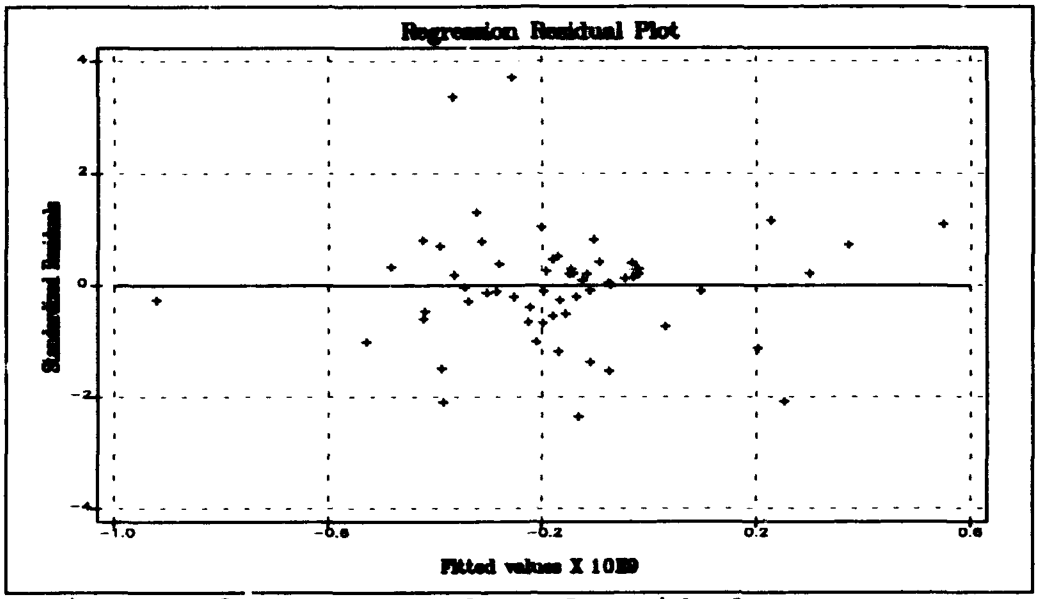


Figure J.3. Scatter plot of residuals versus predicted values of *NPV* for Model 4.

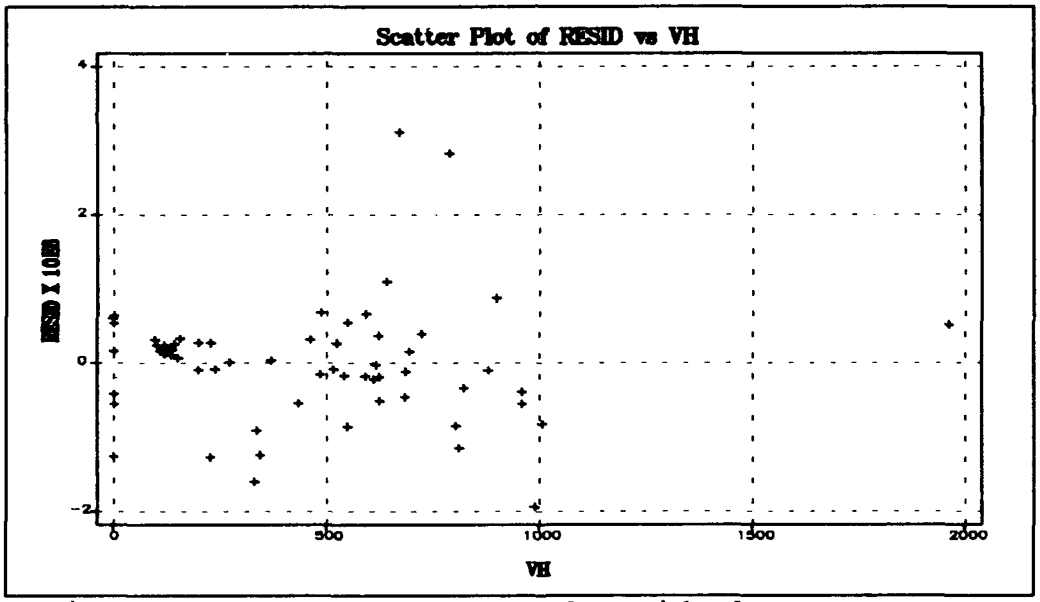


Figure J.4. Scatter plot of residuals versus independent variable *VH* for Model 4.

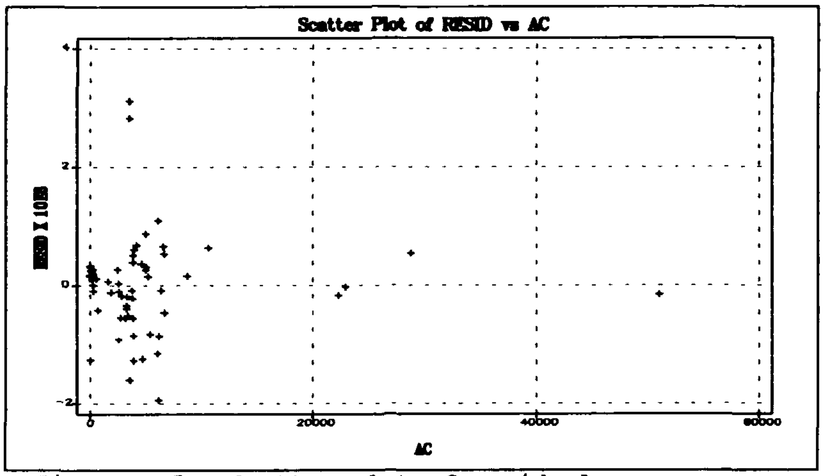


Figure J.5. Scatter plot of residuals versus independent variable *AC* for Model 4.

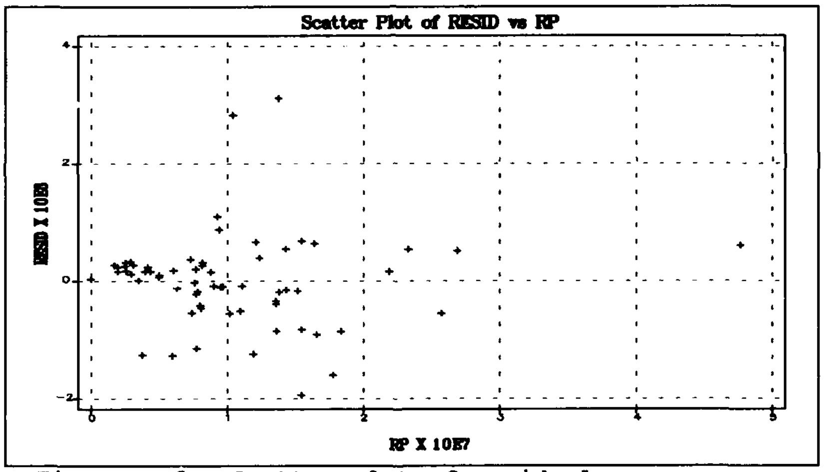


Figure J.6. Scatter plot of residuals versus independent variable *RP* for Model 4.

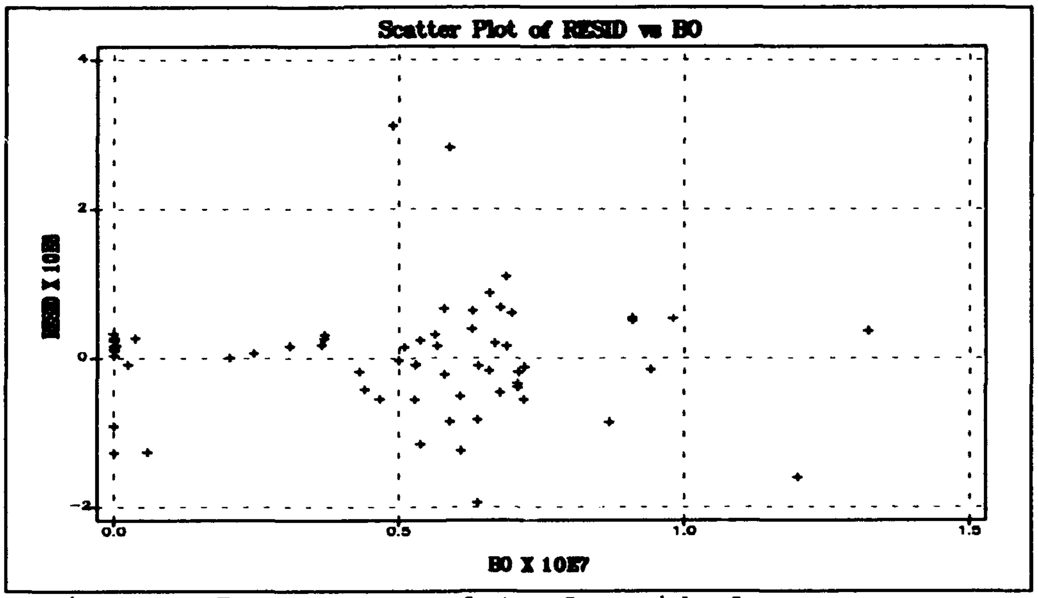


Figure J.7. Scatter plot of residuals versus independent variable *BO* for Model 4.

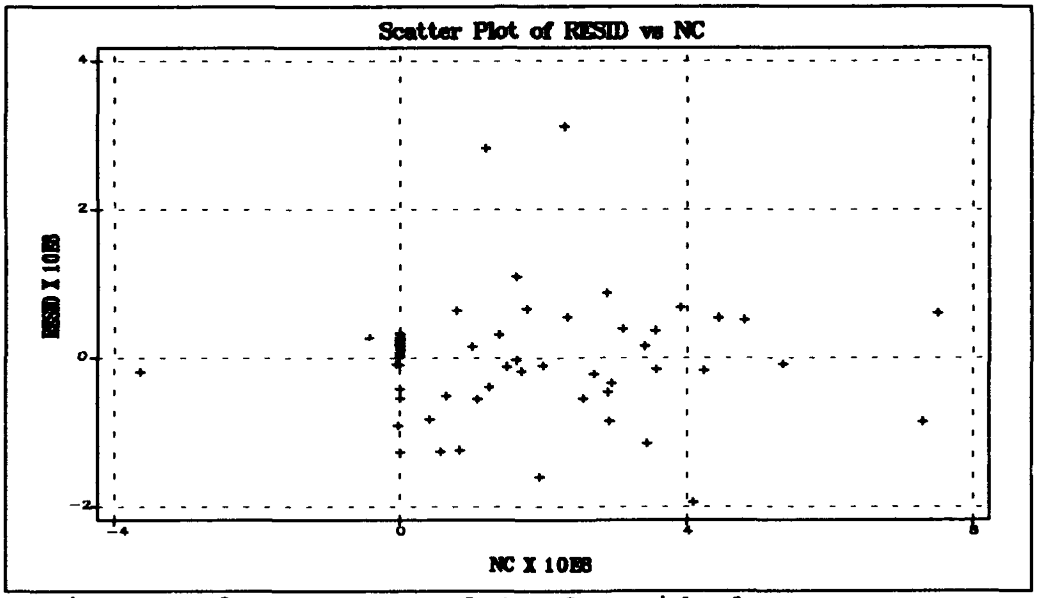


Figure J.8. Scatter plot of residuals versus independent variable *NC* for Model 4.

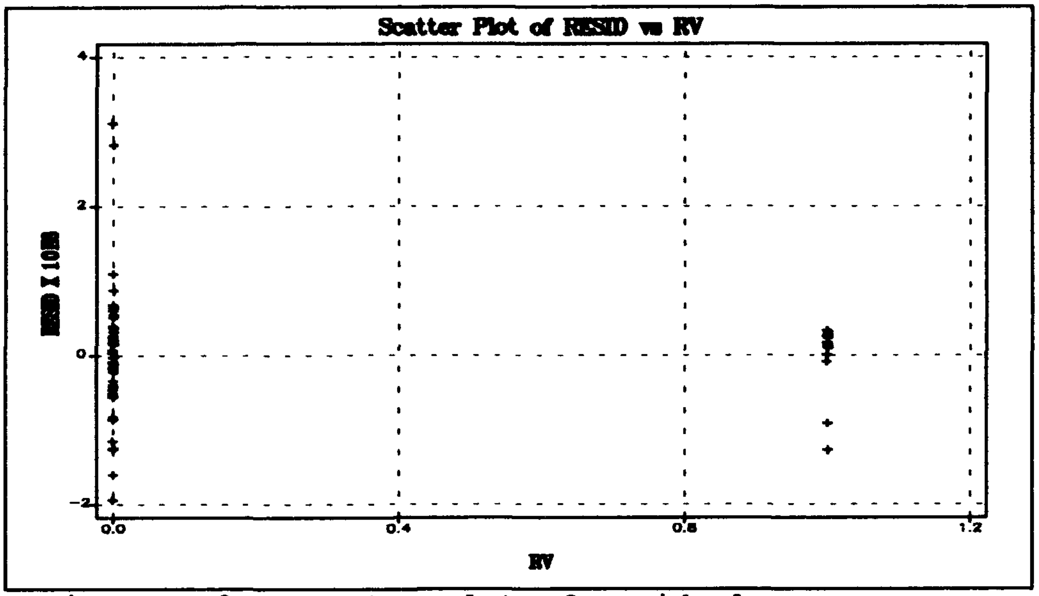


Figure J.9. Scatter plot of residuals versus independent (dummy) variable *RV* for Model 4.

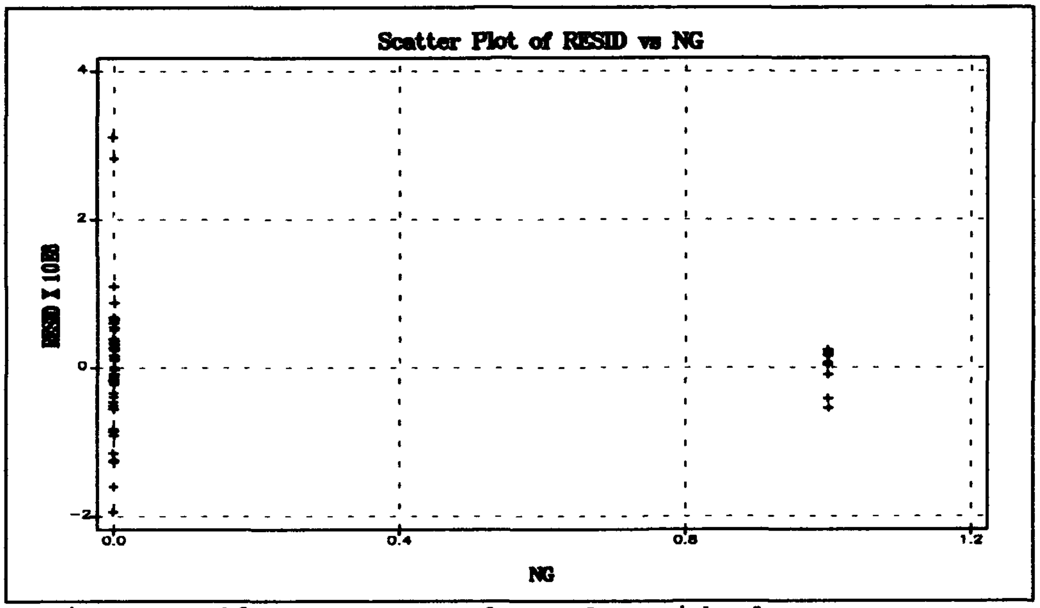


Figure J.10. Scatter plot of residuals versus independent (dummy) variable *NG* for Model 4.

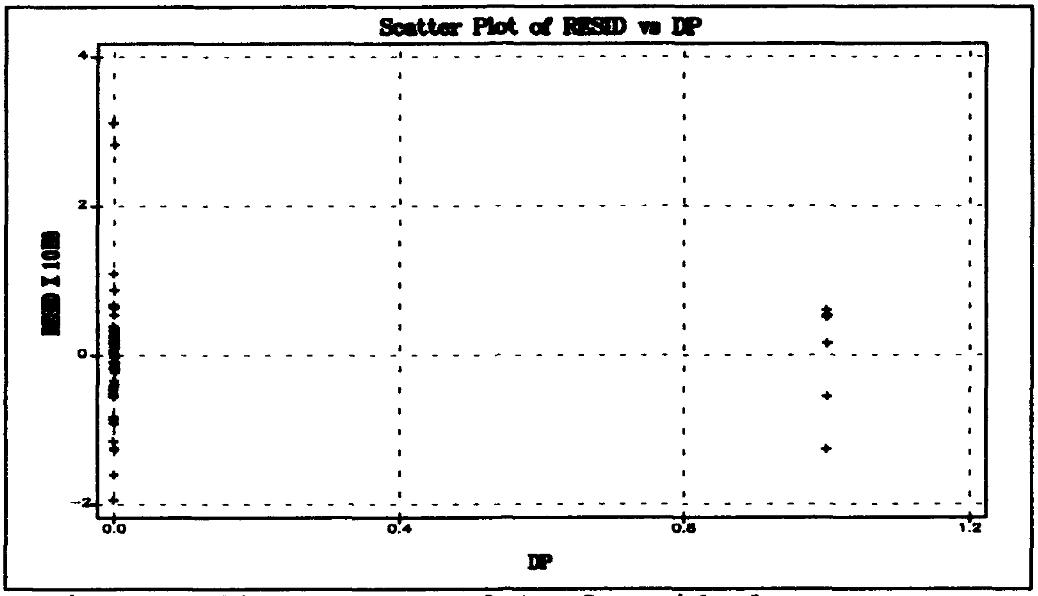


Figure J.11. Scatter plot of residual versus independent (dummy) variable *DP* for Model 4.

<u>Appendix K: Residual Plots for Model 5</u>

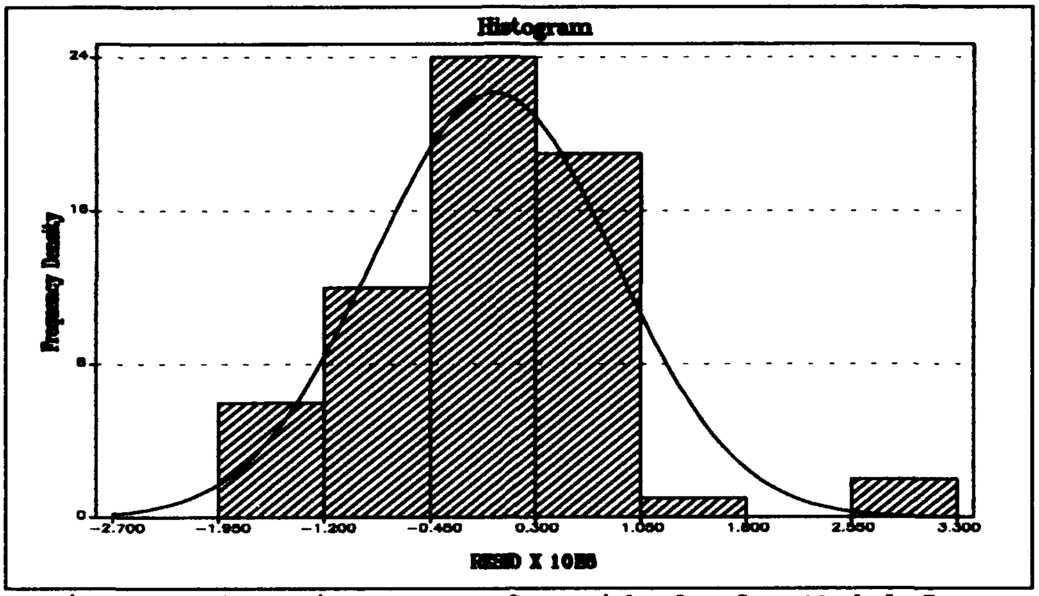


Figure K.1. Histogram of residuals for Model 5.

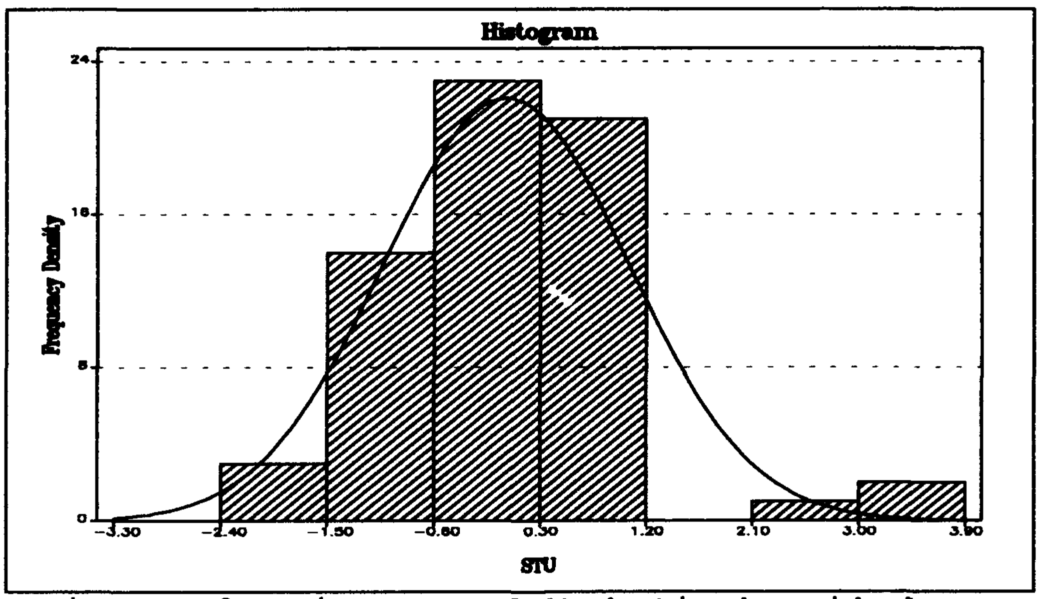


Figure K.2. Histogram of Studentized residuals for Model 5.

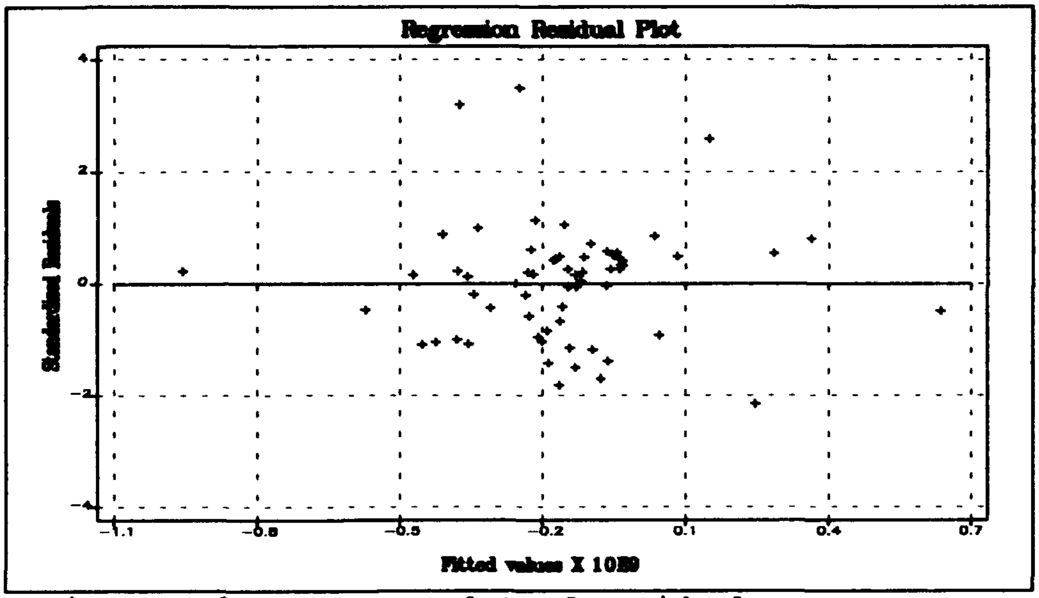


Figure K.3. Scatter plot of residual versus predicted values of *NPV* for Model 5.

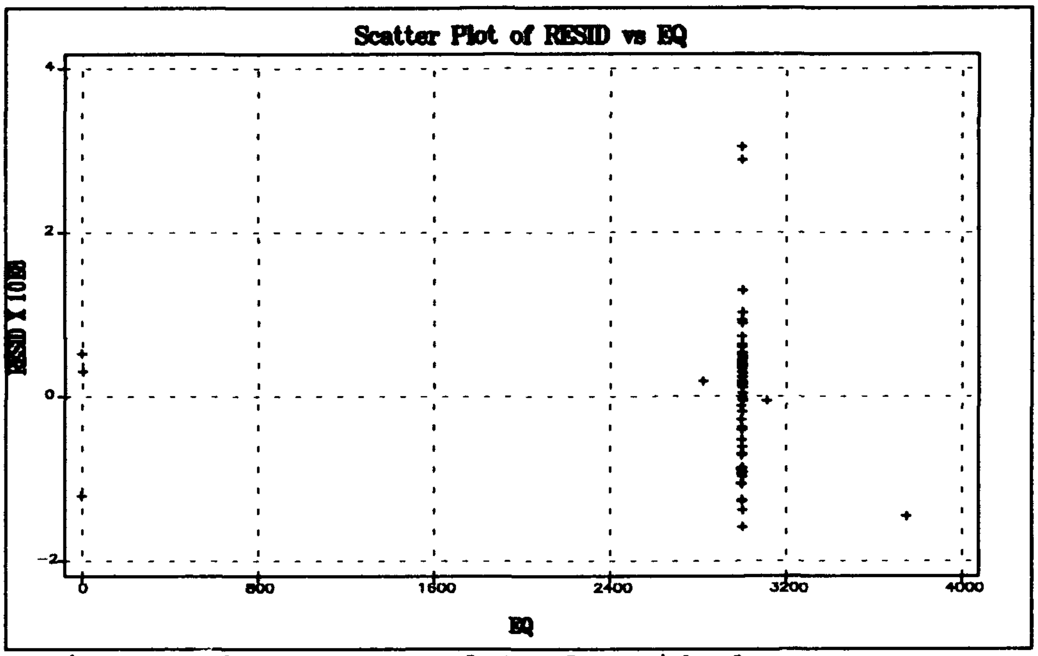


Figure K.4. Scatter plot of residuals versus independent variable *EQ* for Model 5.

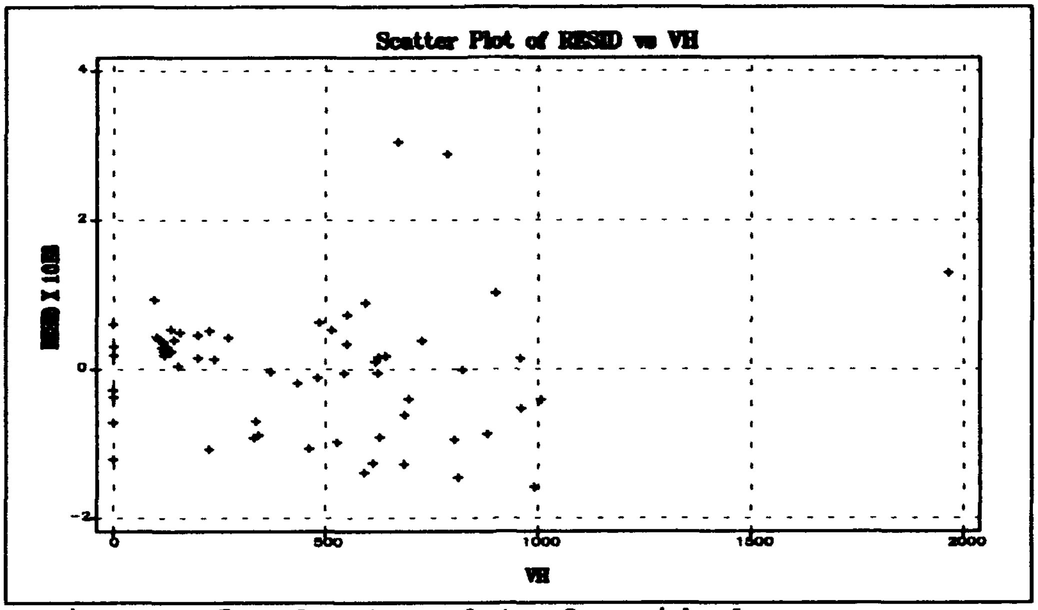


Figure K.5. Scatter plot of residuals versus independent variable VH for Model 5.

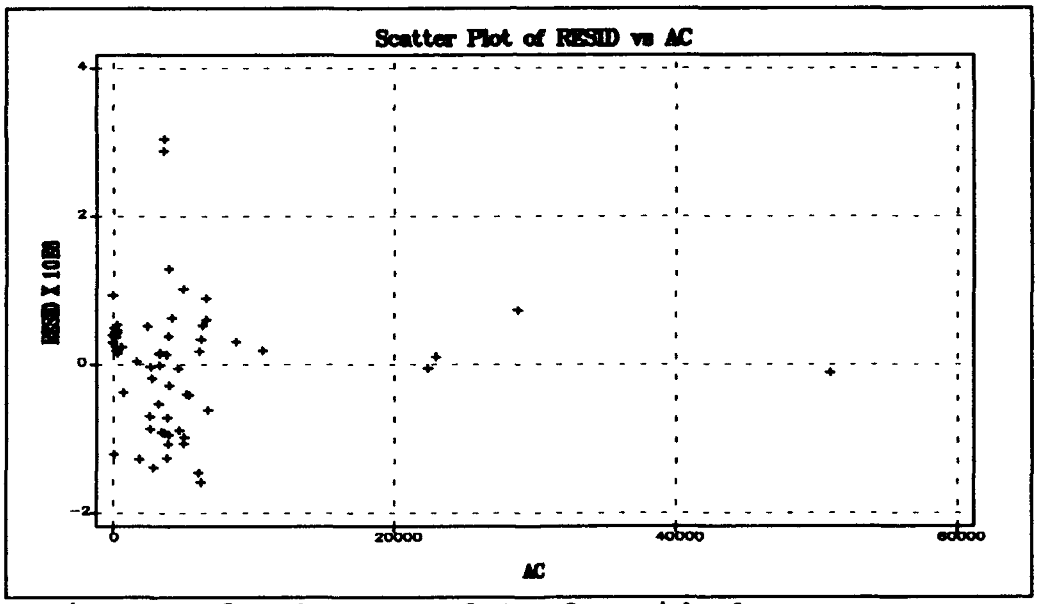


Figure K.6. Scatter plot of residual versus independent variable *AC* for Model 5.

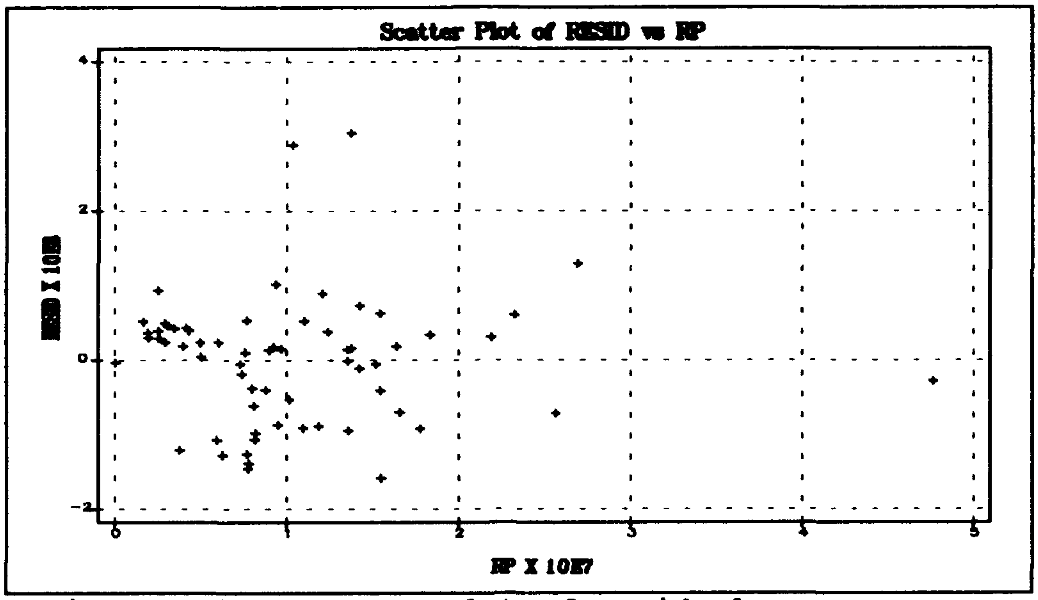


Figure K.7. Scatter plot of residuals versus independent variable *RP* for Model 5.

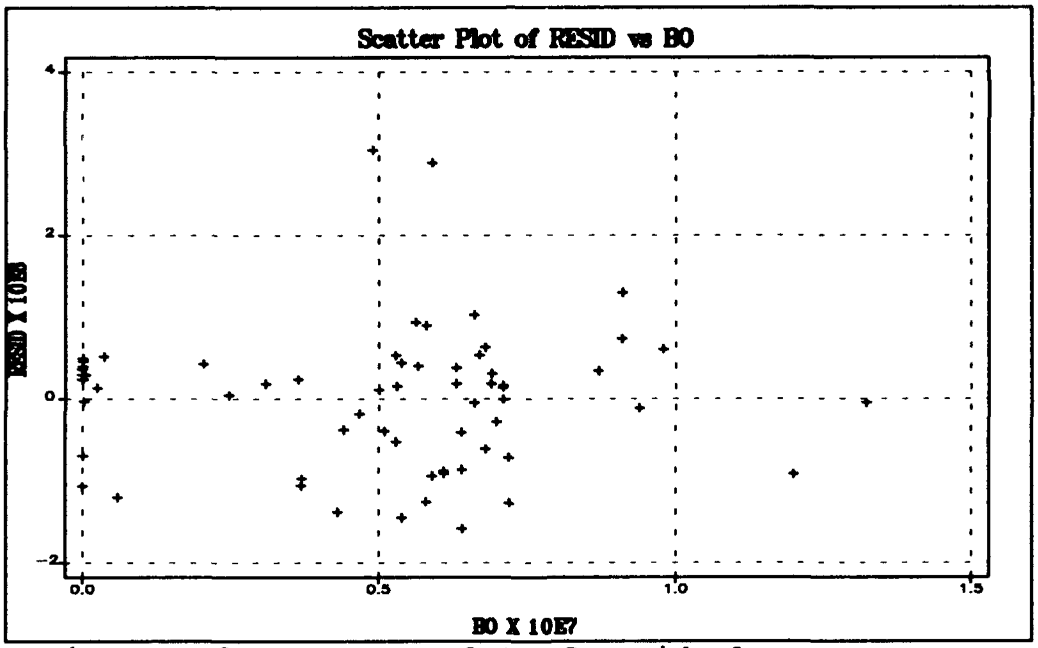


Figure K.8. Scatter plot of residuals versus independent variable *BO* for Model 5.

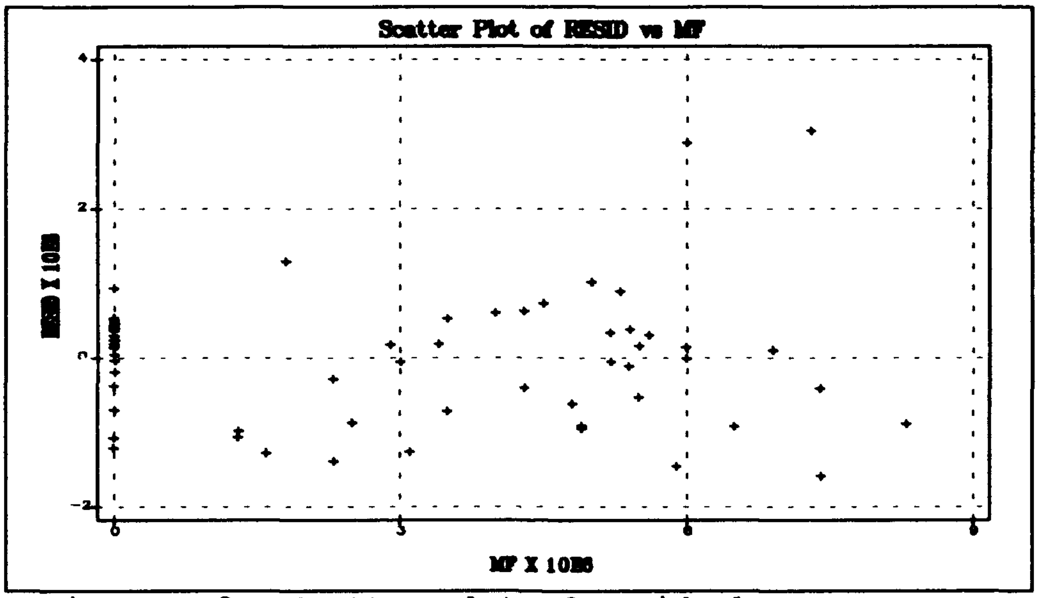


Figure K.9. Scatter plot of residuals versus independent variable *MF* for Model 5.

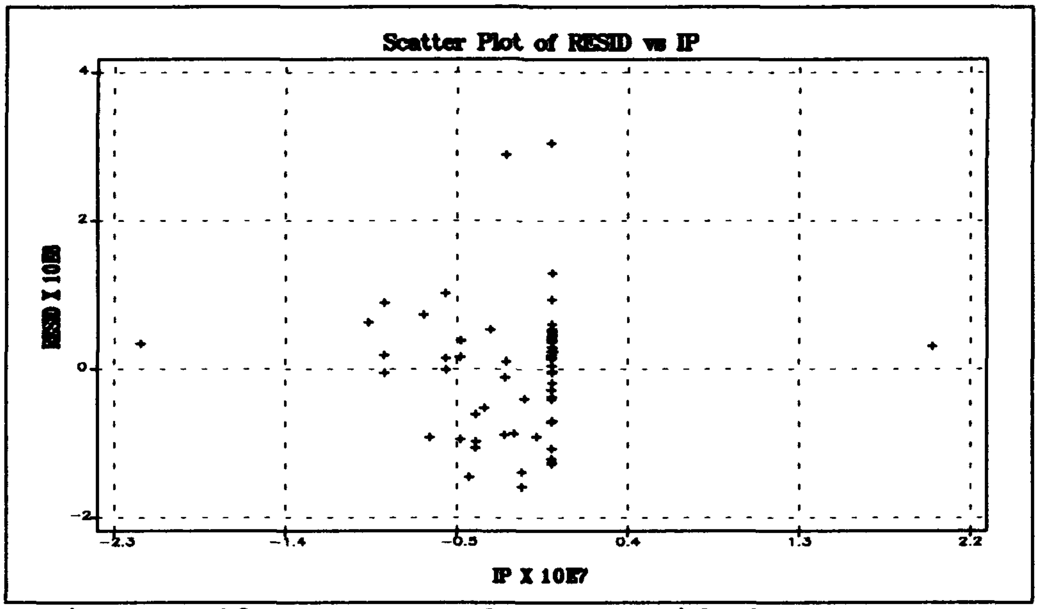


Figure K.10. Scatter plot of residuals versus independent variable for Model 5.

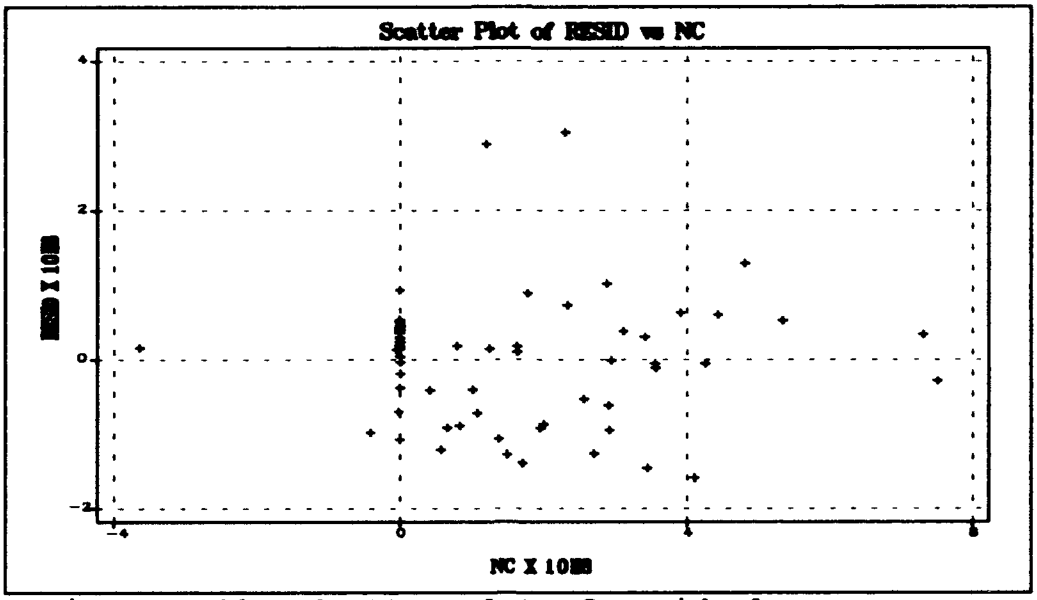


Figure K.11. Scatter plot of residuals versus independent variable *NC* for Model 5.

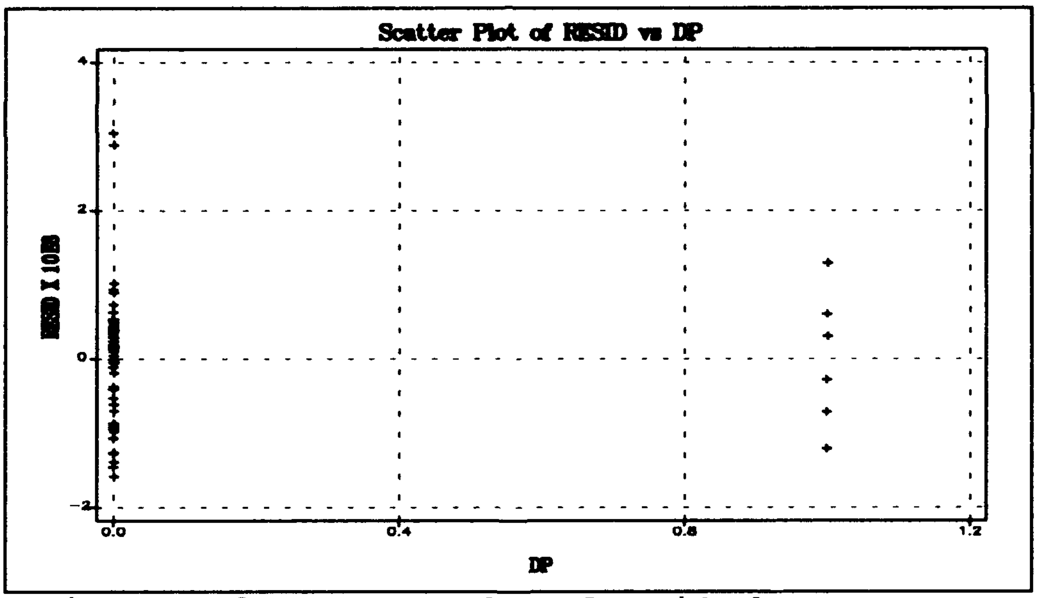


Figure K.12. Scatter plot of residuals versus independent (dummy) variable *DP* for Model 5.

Appendix L: Residual Plots for Model 6

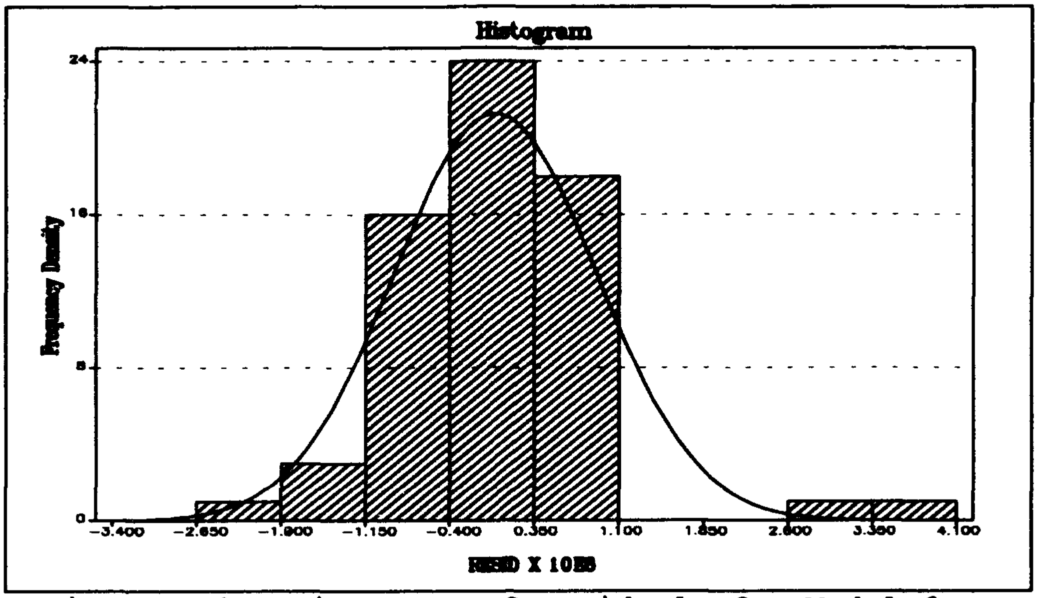


Figure L.1. Histogram of residuals for Model 6.

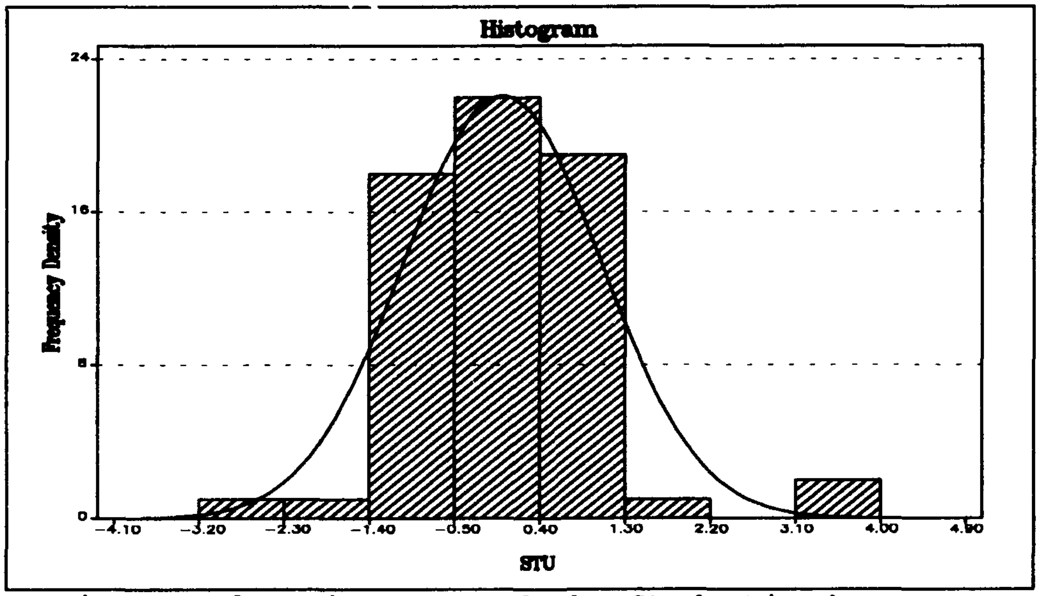


Figure L.2. Histogram of the Studentized residuals for Model 6.

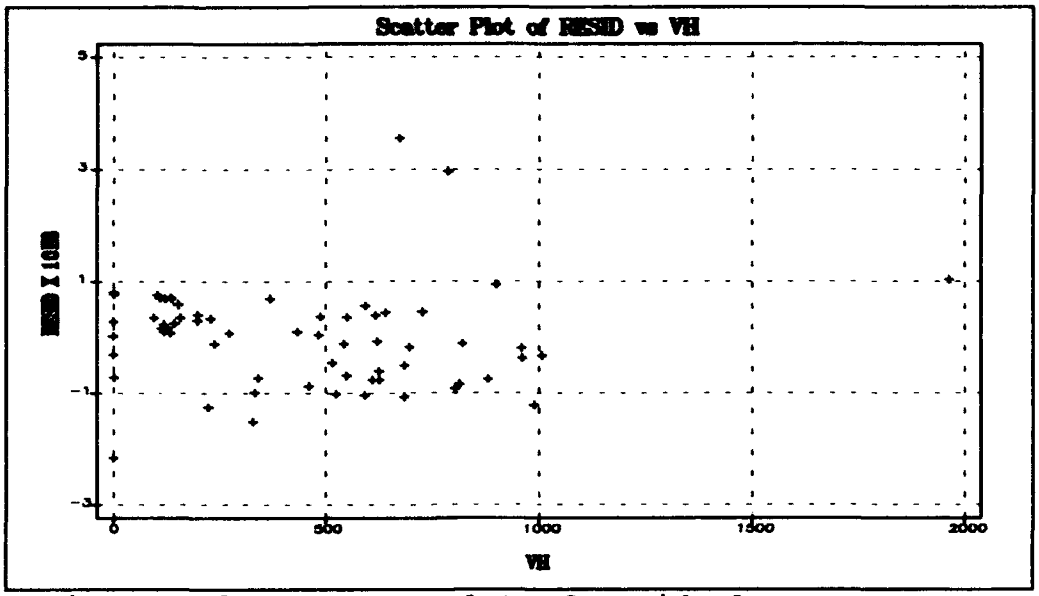


Figure L.3. Scatter plot of residuals versus independent variable *VH* for Model 6.

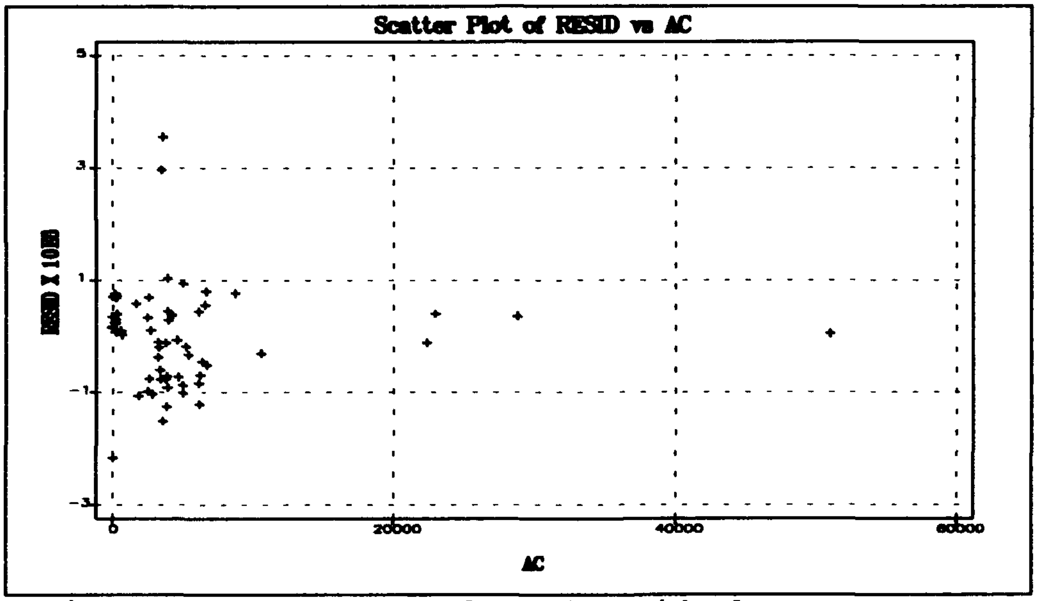


Figure L.4. Scatter plot of residual versus independent variable *AC* for Model 6.

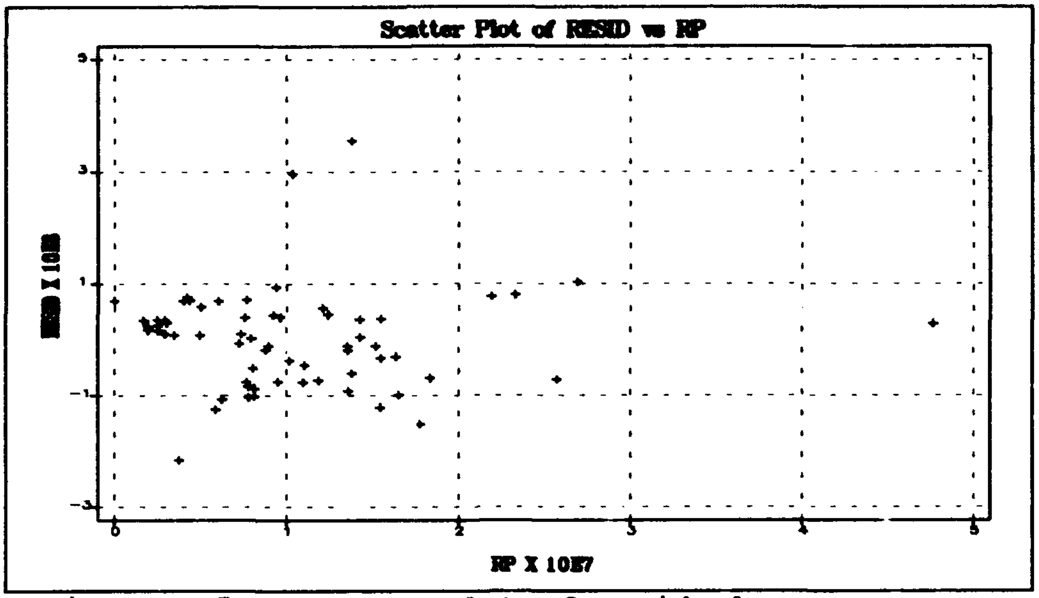


Figure L.5. Scatter plot of residuals versus independent variable *RP* for Model 6.

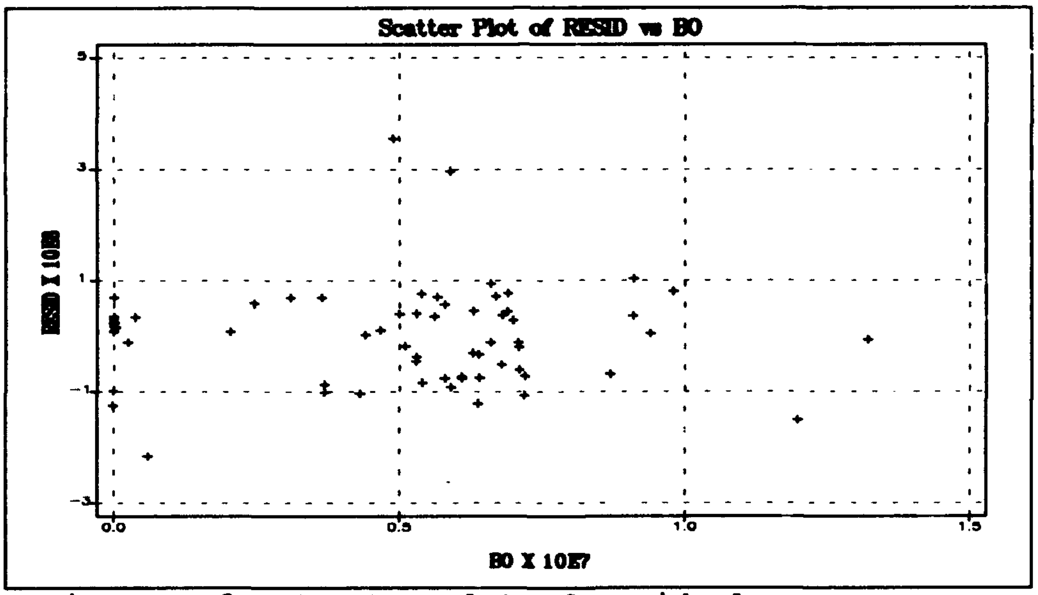


Figure L.6. Scatter plot of residuals versus independent variable *BO* for Model 6.

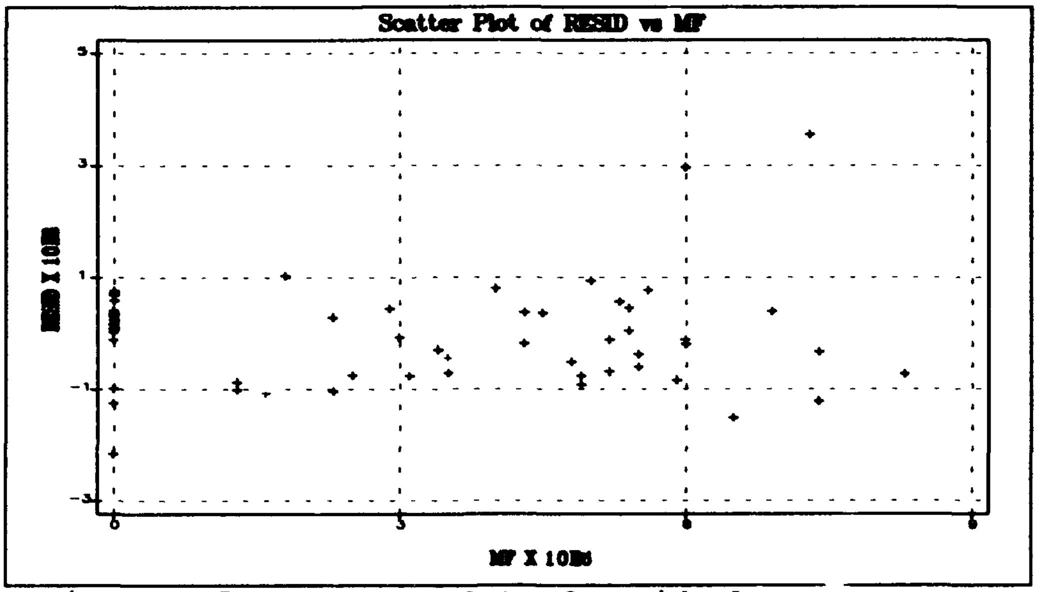


Figure L.7. Scatter plot of residuals versus independent variable *MF* for Model 6.

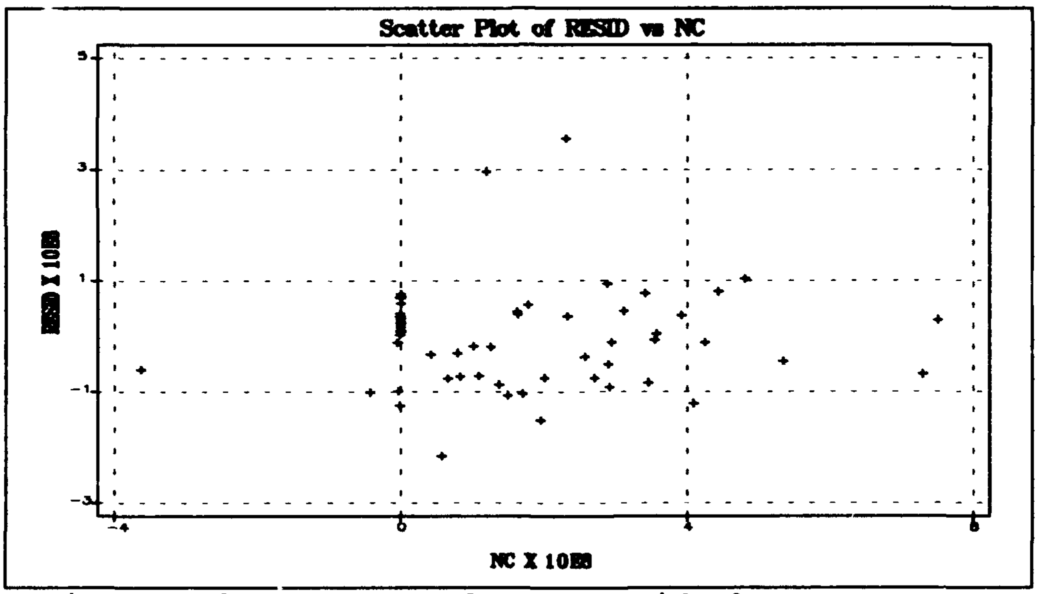


Figure L.8. Scatter plot of residual versus independent variable *NC* for Model 6.

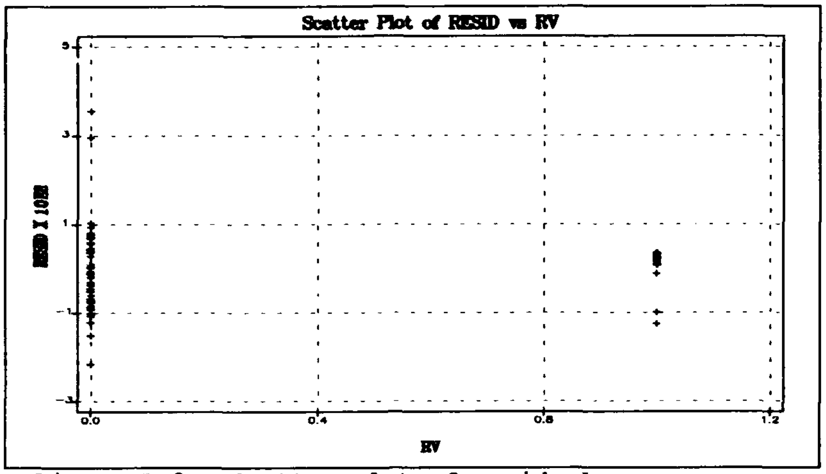


Figure L.9. Scatter plot of residual versus independent (dummy) variable *RV* for Model 6.

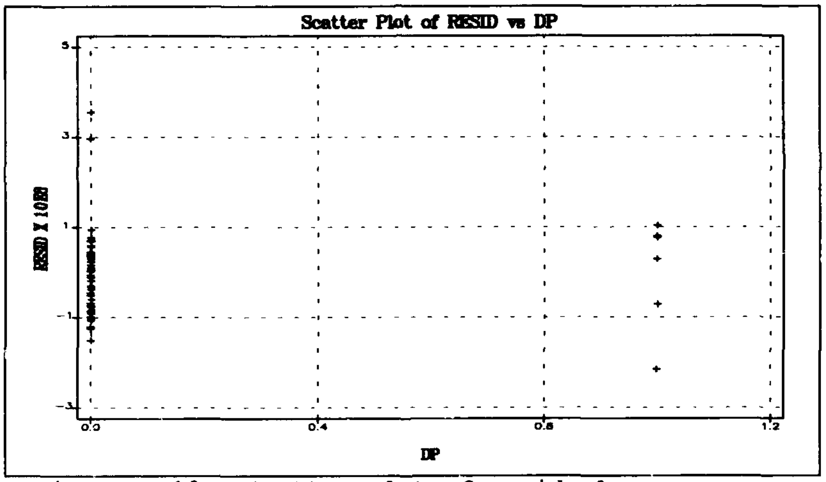


Figure L.10. Scatter plot of residuals versus independent (dummy) variable *DP* for Model 6.

Bibliography

Almand, Major Todd L. An Analysis of The Department of Defense's Environmental Assessment Program. MAS Thesis, ATZL-SWD-GD, U.S. Army Command and General Staff College, Fort Leavenworth KS, June 1991 (AD-B157 692).

Belsley, David A. Conditioning Diagnostics: Collinearity and Weak Data in Regression. New York: John Wiley and Sons, 1991.

Brown, Douglas M. COBRA, The Base Closure Cost Model. Logistics Management Institute. Bethesda, MD (1989).

Cain, J.P. Class lecture, OPER 634, Applied Liner Models. School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB OH, March 1993.

Cody, Ronald P. and Jeffery K. Smith. Applied Statistics and the SAS Programming Language. New York: Elsevier Science Publishing Co., Inc.

Cost of Base Realignment Action (COBRA) Model. Version 4.04, IBM, 1.44M, disk. Computer Software. Richardson and Kirmse Engineering, Inc., 1992.

Defense Secretary's Commission on Base Realignment and Closure. Base Realignment and Closures. Government Printing Office. Washington D.C. (1988)

Department of the Navy. FY 1993 Budget Estimates: DOD Base Realignment and Closure Progam 1. Januarary 1992 (AD-A247404).

Department of the Navy. FY 1993 Budget Estimates: DOD Base Realignment and Closure Progam 1. Januarary 1992 (AD-A247405).

Devore, Jay L. Probability and Statistics for Engineering and the Sciences. Pacific Grove, CA: Brooks/Cole Publishing Co.

Draper, Norman R. and Harry Smith. Applied Regression Analysis, Second Edition. New York: John Wiley and Sons, 1981.

Dupre', First Lieutenant Thomas J. A Cost Prediction Model for Electronic Systems Flight Test Costs. MS Thesis, AFIT/LSSR 108-83, School of Systems and Logistics, Air Force Institute of Technology (AU), Wright Patterson Air Force Base OH, September 1983 (AD-A135598).

Fabrycky, Wolter J. and Blanchard Benjamin S. Life-Cycle Cost and Economic Analysis. New Jersey: Prentice Hall, Inc., 1991.

Fisher, Gene H. Cost Considerations in Systems Analysis. New York: American Elsevier Publishing Company, Inc., 1971.

Gardner, First Lieutenant Thomas E., and Captain Stephen M Passarello. A Parametric Estimating Model for Flight Simulator Acquistion. MS Thesis, AFIT/LSSR 41-81, School of Systems and Logistics, Air Force Institute of Technology (AU), Wright Patterson Air Force Base OH, June 1981 (AD-A103237).

Gatlin, Lt Col Robert B. BRAC, What Will it Cost?. U.S. Army War College. Carlisle Barracks PA, June 1992 (AD-A250842).

"Guide to Air Force Bases" Air Force Magazine, 166-177 (May 1971).

Gunst, Richard F. and Mason, Robert L. Regression Analysis and Its Applications. New York: Marcel Dekker, Inc., 1980.

Holk, Glenn A. The Criteria for and Effects of Base Closures. MS Thesis, Naval PostGraduate School, Monterey CA, December 1989 (AD-A220121).

Kage, Major Gordon D. II. Estimating the Cost of Composite Material Airframes Using the Rand Corporation Development ad Procurement Cost of Aircraft Parametric Model (DAPCA III). MS Thesis, AFIT/LSSR 39-83, School of Systems and Logistics, Air Force Institute of Technology (AU), Wright Patterson Air Force Base OH, September 1983 (AD-A134993).

Langley, Russell. Practical Statistics Simply Explained. New York: Dover Publications, Inc., 1971.

Milton, J.Susan and Arnold, Jesse C. Probability and Statistics in the Engineering and Computing Sciences. McGraw-Hill Book Company, 1986.

Montgomery, Douglas C. and Peck, Elizabeth A. Introduction to Linear Regression Analysis. New York: John Wiley & Sons, Inc., 1982.

Muenchow, Captain William K. Developing a Parametric Cost Model for Estimating the Cost of Buying Out Telephone Switches. MS Thesis, AFIT/GLM/LSQ/91S-48, School of Systems and Logistics, Air Force Institute of Technology (AU). Wright-Patterson Air Force Base OH, September 1991.

Myers, Mryon G., Paul R. McClenon, and William A. Woodring. Cost Estimating Relationship for Real Property Maintenance Activity at Army Installations. Logistics Management Institute. Washington D.C., January 1983 (AD-A126936).

Nelson, Commander Neal J. Spending to Save: Modeling the Closure of NAS Moffet Field. MA Thesis, Naval Post Graduate School, Monterey CA, June 1989 (AD-A216774).

Olver, K.M., John J. Cloos, Daniel B. Levine, and Dennis Madl. Costs and Cost Savings Due to Labratory Realignments. Institute for Defense Analysis. Alexandria, VA. (1991).

Rasher, Donald C. Closing Domestic Military Bases to Generate Cost Savings and Increase Military Efficiency. Strategic Study, NDU-NWC-86-2 A-3S, National Defense University, National War College, Fort McNair Washington D.C., February 1986 (AD-B103440).

SAS Institute, Inc. SAS User's Guide: Basics, Version 5 Edition. Cary NC: SAS Institute Inc., 1985.

-----. SAS User's Guide: Statistics, Version 5 Edition. Cary NC: SAS Institute Inc., 1985.

Sen, Ashish and Srivastava, Muni. Regression Analysis: Theory, Methods, and Applications. New York: Springer-Verlag New York Inc., 1990.

Theil, Henri. Priniciples of Econometrics. New York: John Wiley and Sons, 1971.

Verma, First Lieutenant Rajiv S. The Use of Cost Estimating Relationships Versus Accounting Models For Estimating Maintenance and Repair Costs: A Methodology Demonstration. MS Thesis, AFIT/GSM/LSQ/87S-35, School of Systems and Logistics, Air Force Institute of Technology (AU). Wright-Patterson AFB, OH. September 1987 (AD-A186923).

Wesolowsky, George O. Multiple Regression and Analysis of Variance. New York: John Wiley and Sons, 1976.

Welch, George E. Property Disposal Requirements During Base Closure. MS Thesis, Army Command and Staff College. Fort Leavenworth KS, June 1991 (AD-B160006).

Zimon, Henry A. (USA). The Effects of U.S. Military Base Closures on Regional and Local Economies 1950-1980. Student Report, Air Command and Staff College. Maxwell AFB AL, July 1983 (AD-B074920).

VITA

Captain Douglas D. Hardman was born on 11 February 1966 in Coeur d'Alene Idaho. He is an Eagle Scout with a Bronze Palm Leaf Award. He graduated for Coeur d'Alene Senior High School in 1984 and attended the University of Idaho on an Air Force ROTC Scholarship. He graduated in 1989 with a Bachelor of Science in Electrical Engineering (specialty: High Frequency and Control Systems). He served his first tour of duty at McChord Air Force Base, Washington. He was an Electrical Design Engineer and Chief of Readiness for the 62nd Civil Engineering Squadron. Additionally, he commanded a Prime Base Engineer Emergency Force (BEEF) team in building a Forward Operating Base for the Washington Air National Guard Exercise "Eagle Strike I".

Captain Hardman was then reassigned to Kunsan Air Base, Republic of Korea, in June 1991. He was responsible for the Simplified Acquisition of Base Engineering Requirements (SABER) construction contracts for the 8th Civil Engineering Squadron.

Captain Hardman then entered the Graduate Engineering and Environmental Management program at the Air Force Institute of Technology in May 1992.

Upon graduation, Captain Hardman will be assigned to the Engineering and Services Staff at 5th Air Force at Yokota Air Base, Japan.

He is married to Captain Amy Hardman, a Legal Officer at 5th Air Force at Yokota Air Base, Japan.

Permanent Address: 814 N. 10th
Coeur d'Alene, ID
83814

VITA

Captain Michael Nelson was born in Livingston, Montana on 13 April 1966. He graduated from Hardin Senior High School in Hardin, Montana in 1984. Mike attended Rensselaer Polytechnic Institute on an Air Force ROTC scholarship for two years before transferring to Montana State University. He received a Bachelor of Science degree in civil engineering from MSU in 1989. He reported for active duty to the 23d Civil Engineering Squadron at England Air Force Base, Louisiana, where he held positions as design engineer, Chief of Military Family Housing, and Chief of Readiness Management. This tour was highlighted by his selection as Officer in Charge of the 1990 Readiness Challenge Team.

In May 1992 Captain Nelson was reassigned to Wright-Patterson AFB to pursue a Master of Science degree in Engineering and Environmental Management at AFIT.

Upon graduation from AFIT, Captain Nelson will be assigned to the Environmental Management Directorate of Aeronautical Systems Center at Wright-Patterson AFB.

Mike is married to Captain Kristen Nelson, a manufacturing officer in Aeronautical Systems Center. They have one daughter, Michaela.

Permanent Address: Rt #1, Box 1137
Hardin, Montana
59034

Ohio Address: 6001 Jennagate Lane
Huber Heights, Ohio
45424

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
September 1993

3. REPORT TYPE AND DATES COVERED
Master's Thesis

4. TITLE AND SUBTITLE
A PARAMETRIC REGRESSION OF THE COST OF BASE REALIGNMENT ACTION (COBRA) MODEL

5. FUNDING NUMBERS

6. AUTHOR(S)
Douglas D. Hardman, Captain, USAF
Michael S. Nelson, Captain, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology, WPAFB OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GEE/ENS/93S-03

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
This study develops a parametric model that is capable of generating accurate estimates of the costs to close Air Force installations. The new model is based upon, but much simpler to use than, the Cost of Base Realignment Action (COBRA) model. COBRA is an economic cost analysis model that requires a minimum of 250 inputs and as many as 700 inputs. The new parametric model requires just 10 input variables and was developed using least squares multiple regression. Comparison of the new parametric model to COBRA indicates that it captures 91 percent of the variance in cost estimates generated by the detailed COBRA model.

The 20-year Net Present Value (NPV) of actions to close an installation is the figure of merit used in the new model. The COBRA model and the new parametric model generate similar rank orderings of bases, when NPV is used as the ranking criterion. Use of the Spearman's Correlation Test shows a direct correlation between the rank orders for each model at significance level $\alpha \ll 0.01$.

The parametric model is recommended for use as a precursor to COBRA to narrow the number of contemplated closure installations to just those that require further, more detailed analysis and output.

14. SUBJECT TERMS
COBRA, base closure, cost, cost model, parametric cost model, linear regression, model, realignment

15. NUMBER OF PAGES
182

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL